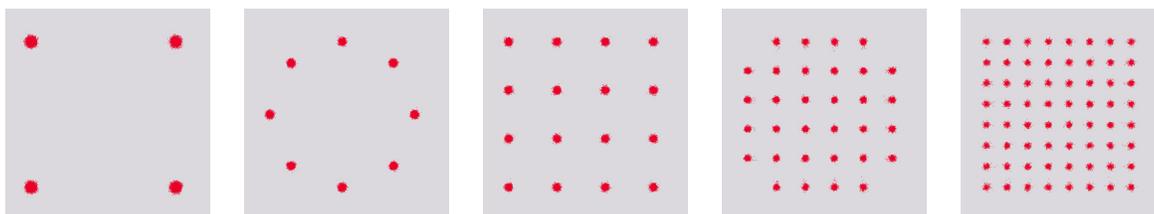


Everything You Need to Know About Complex Optical Modulation



Constellation diagrams for QPSK, 8-PSK, 16-QAM, 32-QAM, and 64-QAM formats

Contents

| | |
|--|-----------|
| Benefits of complex modulation..... | 5 |
| No more limits to spectral efficiency?..... | 7 |
| Shannon–Hartley-theorem..... | 7 |
| Complex Coding Concepts for Increased Optical Bit Transfer Efficiency | 9 |
| How does this happen in practice?..... | 9 |
| Which Modulation Scheme Best Fits My Application? | 13 |
| New speed for the signal | 13 |
| Phase-shift keying schemes | 15 |
| Amplitude- and phase-shift keying schemes | 18 |
| Quadrature Amplitude Modulation (QAM) | 19 |
| Time Domain Pulse Shaping for Increased Spectral Efficiency | 22 |
| Nyquist ISI criterion..... | 22 |
| Nyquist pulse shaping using a finite impulse response filter..... | 23 |
| The concept of raised cosine filters | 25 |
| Raised cosine filters in practice | 27 |
| How much spectral efficiency can be gained? | 28 |
| An Optical Transmitter for Every Need | 31 |
| Controlling the phase with the electro-optic effect | 31 |
| Mach-Zehnder modulator for transmission of a QPSK signal..... | 34 |
| Transmitters for more complex modulation schemes | 35 |
| How to Detect Complex Modulated Optical Signals | 37 |
| How to detect the phase of an optical signal..... | 37 |
| A local oscillator helps | 38 |
| Suppressing phase-independent terms with a balanced receiver..... | 39 |
| Taking the concept to the IQ plane – IQ demodulator..... | 40 |
| Extending the concept to dual polarization..... | 40 |

Contents, Continued

| | |
|--|-----------|
| Use a delayed copy of the signal as a reference – delay line interferometers | 42 |
| Frequency domain detection | 44 |
| Preferences? | 46 |
| Coherent Optical Receivers – the Complete Answer..... | 47 |
| Relieving many impairments | 48 |
| Carrier phase recovery | 48 |
| Finding the Jones-Matrix for recovering the original polarization state..... | 50 |
| Easier in Stokes space..... | 53 |
| Determination of the symbols | 55 |
| Quality Rating of Coherent Measurements..... | 56 |
| Traditional quality parameters | 56 |
| Error vector magnitude (EVM)..... | 58 |
| Signal to noise ratio (SNR)..... | 59 |
| BER estimation..... | 59 |
| Get a bigger picture from the EVM..... | 60 |
| Gain imbalance..... | 60 |
| IQ offset | 61 |
| Quadrature error | 62 |
| Frequency error | 63 |
| IQ magnitude error..... | 63 |
| IQ phase error and laser linewidth | 64 |
| IQ skew..... | 65 |
| XY-polarization skew & imbalance | 66 |
| The Future of Coherent Data Transmission | 67 |

New data centers are being built across the globe, while today's CPUs and RAM ensure latencies so low that it's no problem to map immense amounts of data spread over several servers within a fraction of a second. The more critical question, is whether the rest of the infrastructure can keep pace. Explosively growing amounts of data have become an enormous challenge. To avoid bottlenecks in the near future, the bit-rate efficiency needs to increase at every stage of the data journey.

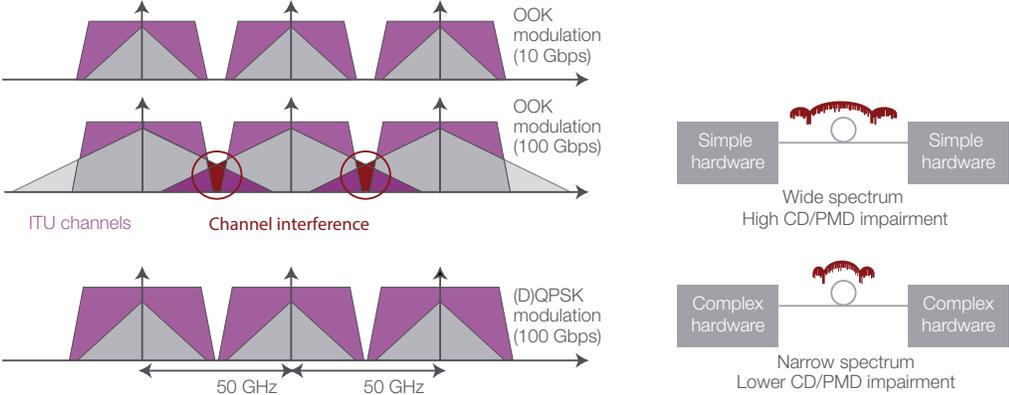


Figure 1. Signal spectrum in ITU grid. With OOK, channel interference or degradation cause serious interference problems at 100 Gbps and beyond; complex modulation schemes can solve these problems

Benefits of complex modulation

Optical data transport started out like its electronic counterpart, with the simplest and therefore cheapest digital coding schemes: return-to-zero (RZ) or non-return-to-zero (NRZ) on/off-keying (OOK). The signal is ideally a rectangular sequence of ones (power on) and zeros (power off). But this concept faced a limitation when transfer rates reached 40 Gbps.

Due to the high clock rate at 40 and 100 Gbps, the bandwidth occupied by the OOK signal becomes larger than the bandwidth of a 50-GHz ITU channel. As can be seen in Figure 1, spectrally broadened channels start to overlap with their neighboring channel and the signals are shaped by the wavelength filters, resulting in crosstalk and degradation of the modulated information.

For this reason, a move away from OOK to more complex modulation schemes, such as differential quadrature phase shift keying (DQPSK), is necessary for high-speed transmission. Complex modulation reduces the required bandwidth, depending on the symbol clock rate and enables higher data rates to be transmitted in the 50-GHz ITU channel plan.

These new concepts also support compensation for chromatic dispersion (CD) and polarization mode dispersion (PMD) via digital signal processing when paired with coherent detection, which provides complete optical field information. Dispersion – an effect caused by the fact that light waves travel at different speeds depending on their frequency and polarization – leads to pulse broadening that degrades the signal if not compensated. Dispersion is especially an issue for long fiber spans. Complex modulation schemes improve spectral efficiency by using all the parameters of a light wave for encoding information: amplitude and frequency or phase. Radio engineers have profited from this approach for many years; now it can be leveraged in the optical world.

The use of coherent detection means that complex optical modulation releases from the need for PMD compensators or dispersion compensating fibers and from the increase in loss and latency these elements induce.

In addition to coherent detection, complex modulation schemes can be combined with other transmission methods to transmit a data signal more efficiently over a fiber link. For example, in polarization division multiplexing (PDM), a second light wave signal, which is polarized orthogonally to the first, carries independent information and is transmitted over the same fiber (see Figure 2). That's like adding a second channel and doubles the transmission speed without the need of a second fiber.

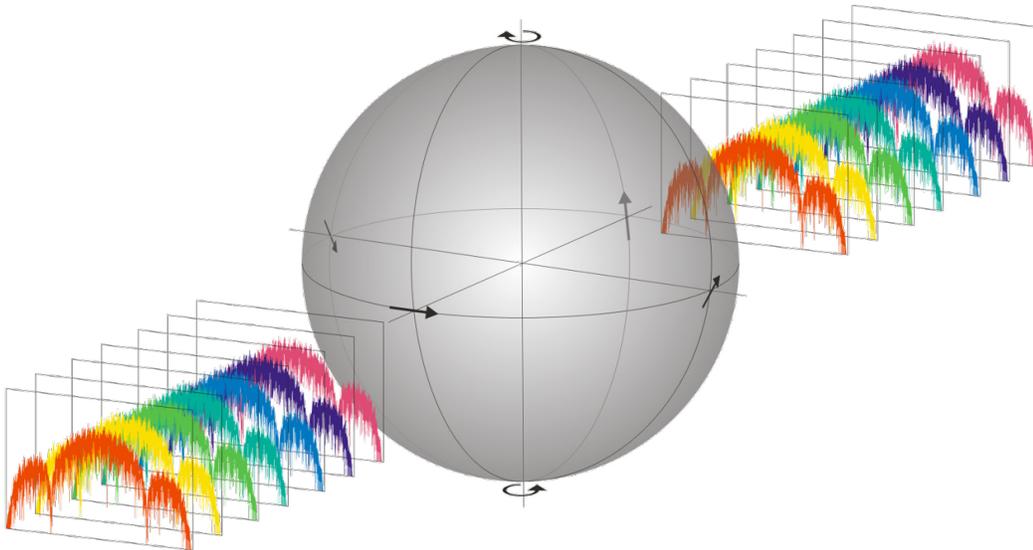


Figure 2. Polarization division multiplexing

Figure 3 gives an idea of how a combination of these different techniques can improve spectral efficiency. At the bottom is the simplest scheme: OOK. Using quadrature phase-shift keying (QPSK) instead, the transfer rate can be doubled while keeping the same symbol rate as in OOK, because in QPSK 2 bits are encoded in one symbol. Another factor of 2 can be gained through PDM. QPSK plus PDM enables the transfer of $2 \times 2 = 4$ times more bits at the same time, meaning at the same clock rate. In the end, after further narrowing the occupied spectrum with a pulse-shaping filter, 100 Gbps can be transmitted in a 50-GHz channel.

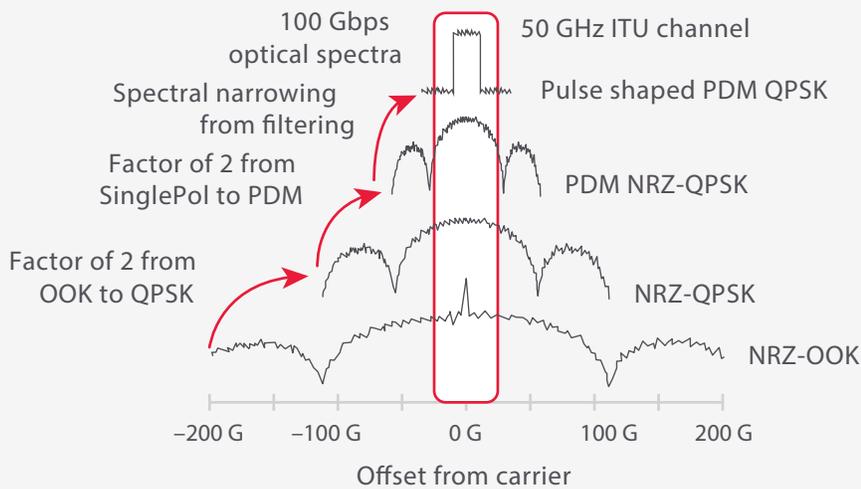


Figure 3. Increasing spectral efficiency for a 100 Gbps data signal by using complex modulation, polarization multiplexing and spectral shaping

Other types of multiplexing (like wavelength division multiplexing (WDM)) continue to be used. The use of pulse shaping filters, which reduce the bandwidth occupied by the signal, completes the tool set.

No more limits to spectral efficiency?

In the 1940s, the American mathematician and electronics engineer Claude Shannon, the 'father of Information Theory', found, that in any communication channel the maximum speed at which data can be transferred without errors can be described in dependence of noise and bandwidth. He called this maximum bit rate 'channel capacity' which is largely known as 'Shannon limit'.

Shannon–Hartley-theorem

Channel capacity:

$$C = B \log_2 \left(1 + \frac{S}{N} \right)$$

Where B is the bandwidth measured in Hertz, S the average received signal power in W and N the average noise power in W.

The channel capacity can be increased by either increasing bandwidth or by optimizing the signal-to-noise-ratio ($SNR = S/N$).

In fact, the theorem provides a theoretical maximum without giving any information about which signal concept allows to get closest to this limit.

In practice, the *SNR* is the fundamentally limiting factor. It is and will also in the future be the subject of ongoing optimization efforts because for data rates beyond 100 Gbps, a better *SNR* performance is needed for long distances to reach the Shannon limit at a given bandwidth.

Ellis, Zhao and Cotter took example parameters to simulate the information spectral density C/B in dependence of transmission and detection type (Figure 4). For non-linear transmission, the information spectral density does not grow infinitely with launch power spectral density. Due to saturation effects of the power amplifier and non-linear effects in the fiber itself, there is a maximum value of information spectral density. This would not be the case if the transmission media were completely linear.

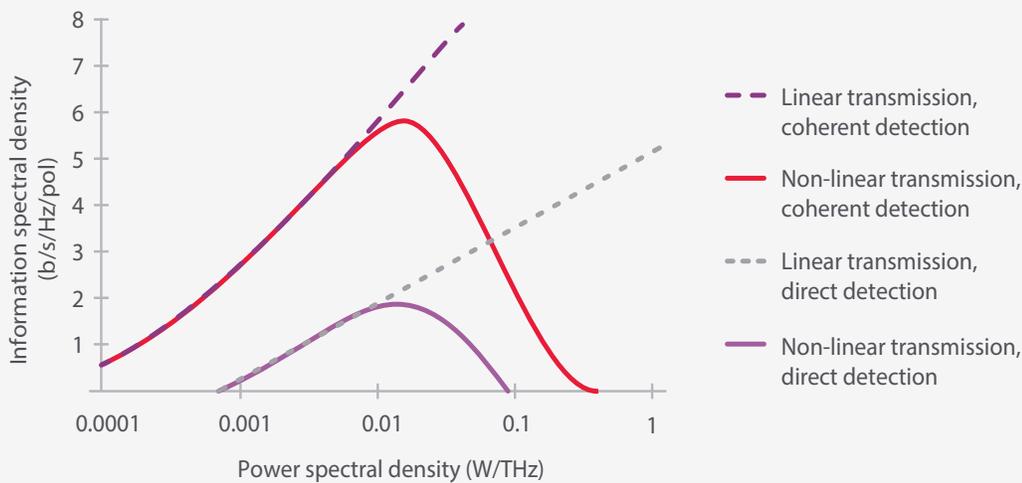


Figure 4. Examples of expected information spectral density limits per polarization¹

This graphic makes it clear that direct detection, used in OOK where information is extracted from amplitude only, cannot compete with coherent detection of complex modulated signals regarding information spectral density.

No doubt that also the different types of complex modulation have a fundamental influence on how close you can get to Shannon’s limit of spectral efficiency. But let’s step back for understanding the basics of coding and modulation schemes first in the next chapter.

1. A.D. Ellis, J. Zhao, and D. Cotter, “Approaching the Non-Linear Shannon Limit,” *Journal of Lightwave Technology*, Vol. 28, No. 4, Feb. 15, 2010.

Complex Coding Concepts for Increased Optical Bit Transfer Efficiency

More efficient techniques than on/off keying (OOK) for transferring bits over the existing fiber infrastructure must be employed to meet the challenges of the cloud revolution and the resulting data avalanche. WDM, for example, has been used very successfully to increase the number of bits transferred over a single fiber and is used for complex modulated signals as well.

Now polarization-division multiplexing (PDM) is being added to double the capacity. In addition to WDM and PDM, it is possible with complex modulation to encode more than one bit within a state of the signal, representing a “symbol”.

This is where complex transmission demonstrates its huge potential. Instead of transmitting a binary data stream, several bits can be coded to a new symbol, and a stream of these symbols can then be transmitted. Figure 5 illustrates this for 2 bits being coded to one new symbol. In this way, twice the amount of data can be accommodated in the same bandwidth.

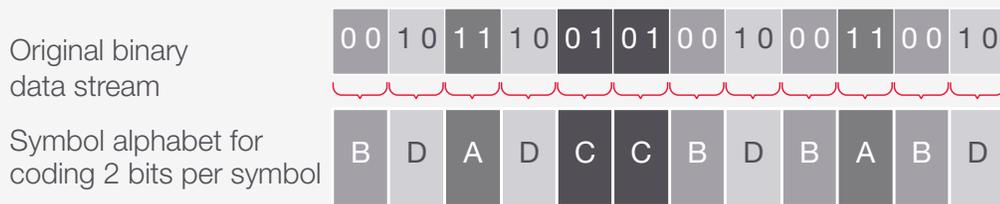


Figure 5. Coding concept: use of symbols to represent a series of bits; here two bits are represented by one alphabetic symbol

Of course, with this knowledge one can imagine schemes where a much larger number of bits are defined by a single symbol that enables a data rate many times greater than in OOK, where a series of ones and zeros is transmitted.

How does this happen in practice?

In OOK, the approach is basically that when the laser source is turned on, this is interpreted as a “one,” and when it is turned off, this reflects a “zero.” In other terms, when the light amplitude exceeds a certain level, this is a one, and a zero is when the amplitude falls below this level.

However, as a light wave is defined by more parameters than just amplitude, more possibilities to encode information are available when using all of the light wave’s degrees of freedom.

Figure 6 shows the mathematical description of the electric field of an electromagnetic wave with two polarization components E_x and E_y . These orthogonal components are used in PDM as two different channels to transfer independent signals. In WDM, different frequencies ω are applied as different channels for independent data transfer at these frequencies/wavelengths. For complex modulation schemes now, additionally to the amplitude E , the phase ϕ of a light wave is modulated for defining the above described symbols.

$$\vec{E} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} E_x e^{i\phi_x} \\ E_y e^{i\phi_y} \end{bmatrix} e^{i(\omega t - k \cdot z)} = \begin{bmatrix} I_x + iQ_x \\ I_y + iQ_y \end{bmatrix} e^{i(\omega t - k \cdot z)}$$

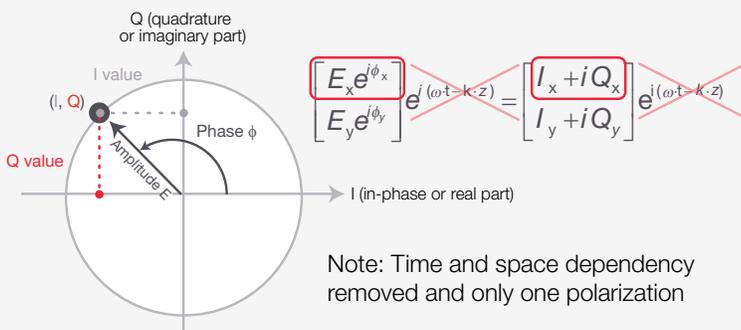
- Polarization division multiplexing
- Amplitude modulation
- Phase modulation
- Frequency (wavelength) division multiplexing

Light is a transversal electromagnetic wave

Use all degrees of freedom to encode information.

Figure 6. Mathematical description of an electromagnetic wave (electric field)

The electric field of the modulated light wave can also be described in the complex plane with an I/Q diagram. Here, I is the in-phase or real part and Q the quadrature or imaginary part as shown in Figure 7 (after removal of time and space dependency of the wave and for one polarization plane only). A symbol corresponds to a point (also called a “constellation point”) in this diagram (which itself also is referred to as a “constellation diagram”) and is defined by a I and an Q value or in polar coordinates by amplitude E and phase ϕ . The constellation points correspond to the symbol clock times and are also called “detection decision points.”



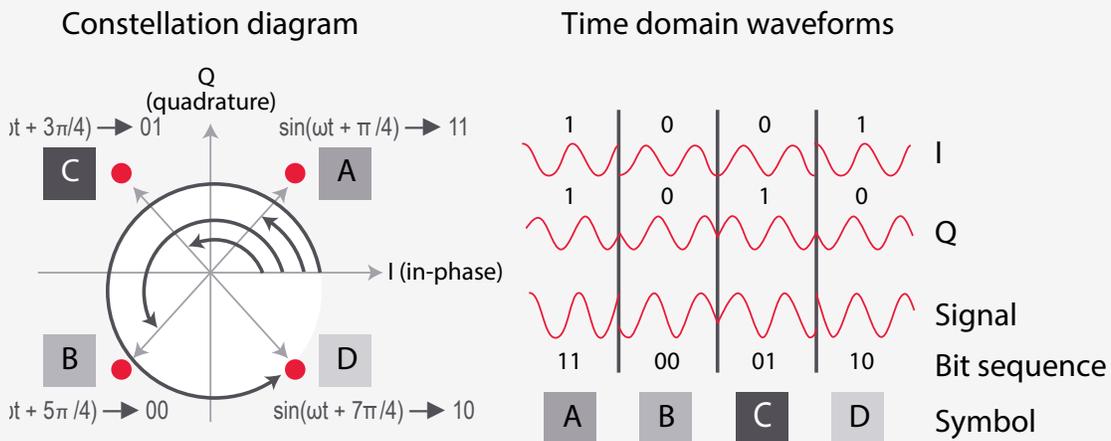
Every carrier signal can be described by two parameters:

- Amplitude
- Phase

Both parameters can be modulated to carry information

Figure 7. I/Q representation of a symbol

Figure 8 shows the constellation points for the four symbols in quadrature phase-shift keying (QPSK), a complex modulation type where the four symbols encode 2 bits each. The constellation points are situated on a circle with radius E . This means that the symbols only differ in phase (always $\pi/2$ between the neighboring points), not in amplitude. In the time domain, the four symbols are reflected by a combination of two waves of the same amplitude and different phase.



There are 4 possible vectors.

One vector position in the complex plane codes 2 bits

Figure 8. Four symbols/constellation points for 2 bits encoded in one symbol (here quadrature phase shift keying (QPSK))

Conventional OOK can also be represented by a constellation diagram. As information is in amplitude only, the bit value 1 can be anywhere on a circle with radius (= amplitude) E (see Figure 9).

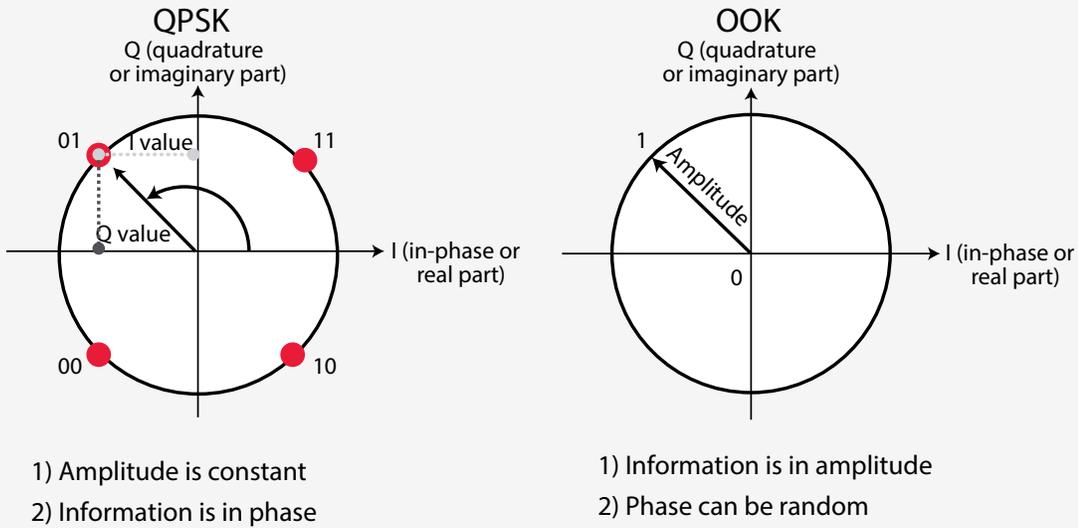


Figure 9. Constellation diagram of QPSK modulation versus OOK; in OOK phase is random

Which Modulation Scheme Best Fits My Application?

This chapter explains how complex coding reduces the required optical bandwidth. In addition, several modulation schemes will be introduced and compared to provide guidance on the ideal scheme for a selected speed class and reach.

New speed for the signal

Now there are effectively two different speeds. First is the bit rate f_{tx} , measured in bits per second, also referred to as “transmission rate.” Second, the symbol rate S quantifies the number of symbols transmitted per second measured in baud. Therefore, it is often called “baud rate.” With $N_{symbols}$ as the number of symbols in the alphabet, the symbol rate calculates as:

$$S = \frac{f_{tx}}{\log_2(N_{symbols})}$$

Figure 10 illustrates this formula for QPSK. If the signal is also polarization multiplexed, divide this result by 2. For a 100-Gbps QPSK signal, this means, for example, that the symbol rate $S = (100 \text{ Gbps}) / (2 \text{ bits/symbol}) / (2 \text{ polarizations}) = 25 \text{ Gbaud}$. The minimum occupied optical bandwidth is then 25 GHz.

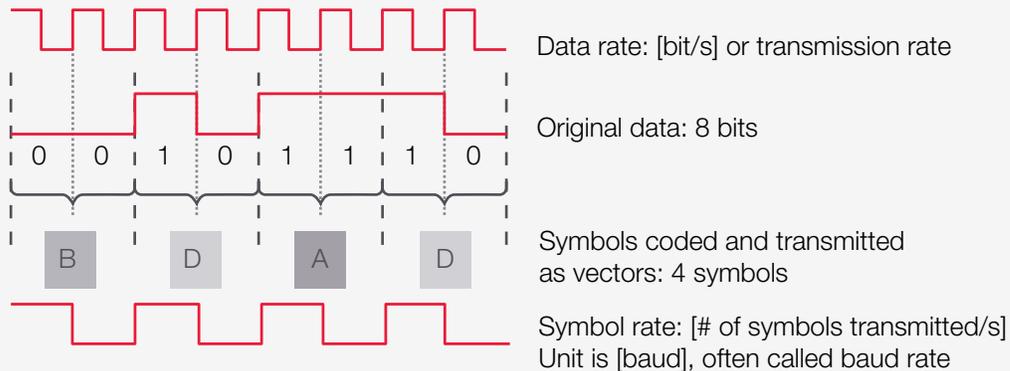


Figure 10. Symbol rate versus data rate for QPSK

Therefore, the bandwidth required by complex modulated signals doesn't depend on the data rate but only on the symbol rate. That also means the more bits encoded into one symbol at a given data rate, the greater the reduction in occupied optical bandwidth. Or in other words, when keeping the symbol rate constant, the data rate can be increased by increasing the number of bits per symbols while keeping the occupied optical bandwidth constant (Figure 11). This technique is commonly used to increase spectral efficiency.

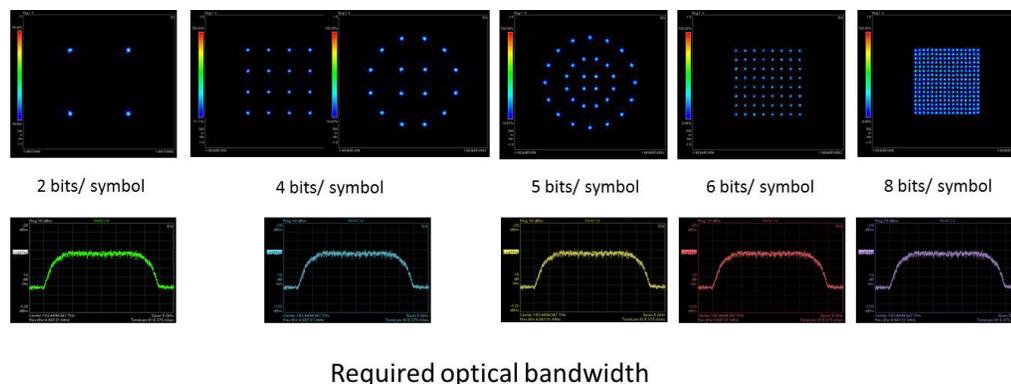


Figure 11. The data rate can be increased by increasing the number of bits encoded in one symbol. The required optical bandwidth stays the same for constant symbol rate

It would be taking too narrow a view though to conclude that in every case a modulation scheme with a higher number of bits encoded into one symbol is the right choice. Apart from the occupied bandwidth, the reach, technical feasibility, existing infrastructure, etc., have to be considered. Modern DSP ASICs often support various data rates through configurable modulation formats with different number of bits per symbol and optimize thereby the spectral efficiency for respective throughput and reach requirements.

Phase-shift keying schemes

In phase-shift keying schemes, the amplitude is constant and information lies in the phase only. Traditional techniques like WDM and polarization-division multiplexing (PDM) can always be used together with any phase-shift keying coding schemes to gain additional data transfer capacity. Multiplexing defines several channels where different phase-modulated signals can be transmitted (Figure 12).

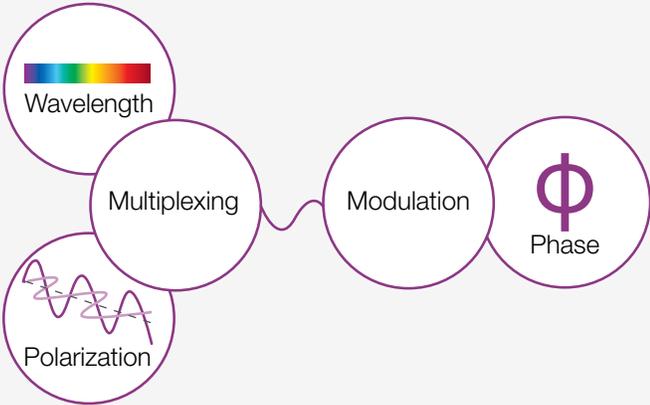


Figure 12. Different ways of gaining data transfer capacity

Binary phase-shift keying (BPSK). BPSK is the simplest pure phase-shift keying format. There are two possible phase values separated by 180° , for example: 0 and π (see Figure 13). The amplitude is constant. The large distance between the two symbols makes it fairly immune against distortions and noise compared to on/off keying (OOK). That makes BPSK suitable for ultra-long-haul applications like submarine fiber networks at speeds up to 40 Gbps.

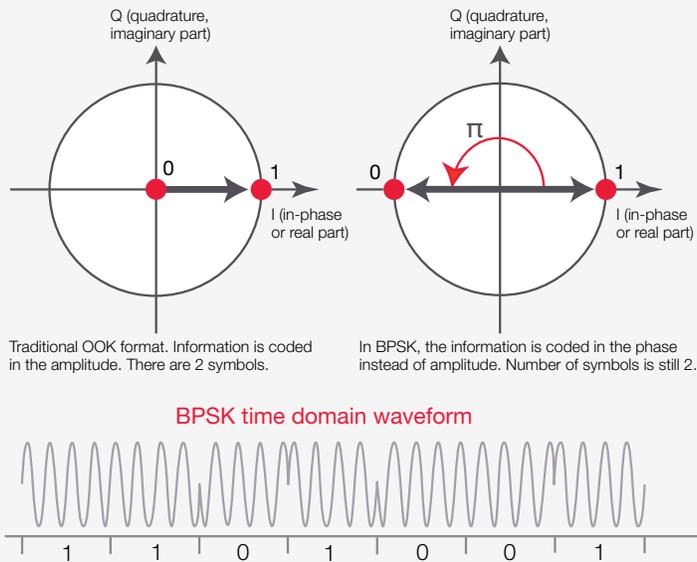


Figure 13. BPSK

The downside of BPSK, however, is that each symbol represents only 1 bit, just like OOK, which makes it equally inapplicable to 100 Gbps and higher data-rate applications. A disadvantage of BPSK versus OOK is there's no simple solution for determining the absolute phase of a signal. More complex and therefore expensive methods are needed; so-called coherent detection is the key term here. In the case of OOK, the amplitude carrying the information can be detected with just a photodiode. That's also referred to as "direct detection."

Differential phase-shift keying (DPSK). To avoid the necessity of coherent detection, BPSK can be modified. In a DPSK signal, not the absolute phase π but a phase change by π reflects a zero. If the phase does not change from one bit to the next, this is interpreted as a 1. On the receiver side, the data stream is split into two identical streams that are then delayed by one bit period. Mixing the two data streams (see Figure 14) results in a signal that can be directly detected by a photodetector. The mixing also offers the advantage of a gain in intensity.

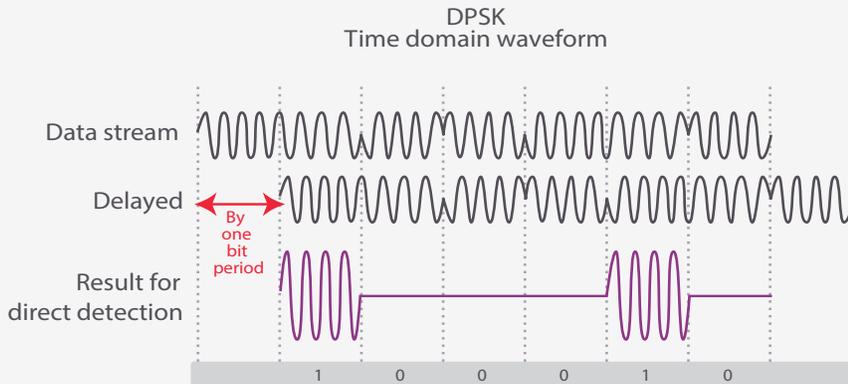


Figure 14. DPSK signals become directly detectable

But just like in OOK, there is only 1 bit transferred per symbol, which makes DPSK inapplicable for the highest data-rate applications. It is a good choice though for speeds up to 40 Gbps for long-haul and ultra-long-haul.

Quadrature phase-shift keying (QPSK). As shown in the previous chapter, one symbol represents 2 bits in QPSK. The four constellation points differ in phase by $\pi/2$, and the amplitude is constant (see Figure 8). In comparison to OOK and BPSK, QPSK enables the data rate to be doubled while keeping the same bandwidth, meaning it's also possible to stay at the same data rate at half the bandwidth. QPSK is commonly used for data rates of 100 Gbps. Coherent techniques are needed for detecting the phase of the signal.

Differential quadrature phase-shift keying (DQPSK). To avoid phase ambiguity in QPSK at the receiver side due to phase shifts induced by the fiber, just like in BPSK, a differential variation of QPSK can be used: DQPSK. The information again is in the phase shifts between the transmitted symbols. The four possible phase-shift values are commonly: 0, $-\pi/2$, $\pi/2$, or π .

In comparison to OOK and BPSK, the data rate can be doubled while keeping the same bandwidth. Alternatively, it is possible to stay at the same data rate at half the bandwidth.

DQPSK shows lower BER than QPSK, but it's less tolerant to dispersion. As in OOK and DPSK, a DQPSK signal can be detected directly.

Amplitude- and phase-shift keying schemes

In amplitude- and phase-shift keying schemes, information lies not only in phase, but also in amplitude. Traditional techniques like multiplexing can always be applied to further increase the amount of data transferred per unit time (Figure 15).

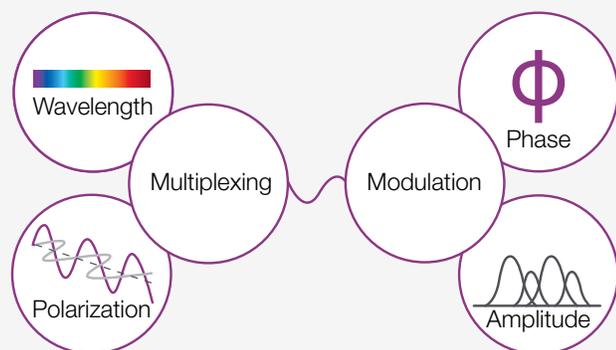


Figure 15. Amplitude-phase modulation schemes with multiplexing

Quadrature Amplitude Modulation (QAM)

With data rates reaching 400 Gbps and higher, this is accomplished using QAM modulation schemes. The modulation and demodulation of QAM signals are more complex and in turn more expensive than alternative formats. On the other hand, the constellation points of higher-order QAM lie further apart than in pure PSK schemes like BPSK or 8-PSK. That makes them less susceptible to noise and distortions, which results in a lower BER.

In a 2^n -QAM scheme, the 2^n -constellation points represent a series of n bits each, usually distributed in a square lattice (see Figure 16). The lowest order QAM, 2-QAM, encodes just 1 bit per symbol. The amplitude is constant and there's a phase difference of π between the two constellation points that correspond to 1 and 0. So 2-QAM is really the same scheme as BPSK. Similarly, the concept of 4-QAM may be different than QPSK, but the resulting constellation diagram is the same. Here as well, there is only one amplitude and the phase of the four constellation points differs by $\pi/2$. In 8-QAM, there are two possible amplitudes and four phases differing by $\pi/2$ that define the constellation points representing a series of 3 bits each.

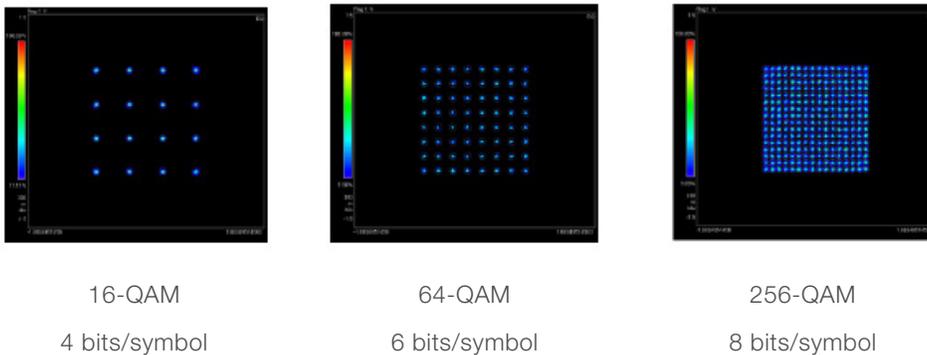


Figure 16. Constellation points of QAM-schemes distributed in a square lattice

As in all other 2^n -QAM schemes with n being an odd number, it's difficult to distribute the constellation points in a square lattice (compare in Figure 17). That has a negative impact on the BER performance; therefore, these 8-QAM schemes play a minor role in practice. Instead, usually 16-QAM is preferred for its double spectral efficiency at only slightly increased BER.

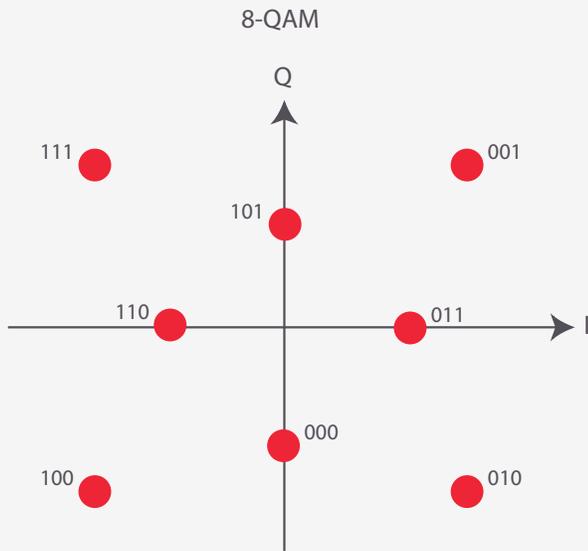


Figure 17. 8-QAM constellation diagram

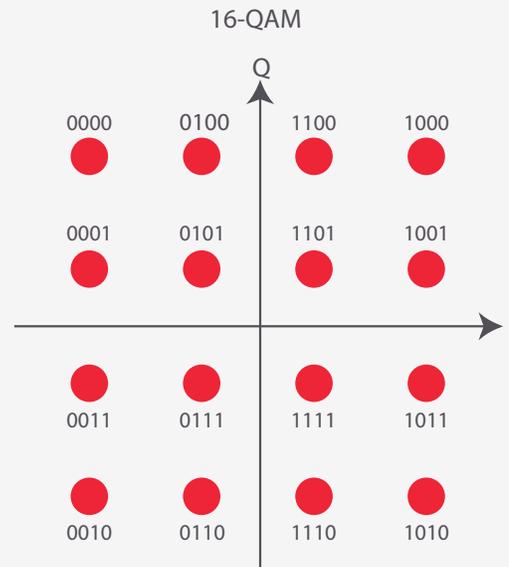


Figure 18. 16-QAM constellation diagram

16-QAM. In this scheme, 4 bits are represented by one symbol. The 16 constellation points are distributed in a square lattice (see Figure 18). Typically, they are Gray coded; from one constellation point to every neighboring point, only one bit value changes. This way, if due to noise, a measured point is wrongly assigned to a neighboring point, the resulting bit error is kept to a minimum: 1 wrong bit.

Amplitude- and phase-shift keying. In APSK, as its name implies, both amplitude and phase are modulated. It differs from QAM in that the constellation points are distributed on concentric circles in the I/Q-plane.

The concept was introduced for satellite systems where the RF power amplifiers show nonlinear behavior. Thus a scheme tolerant to nonlinear amplification was needed—a concept with fewer amplitude states—so that nonlinearities can be balanced more easily.

Comparing the constellation diagrams of 16-QAM and 16-APSK in Figure 19, in 16-QAM there are three amplitudes whereas 16-APSK has only two. In 32-QAM then, there are five amplitudes versus three in 32-APSK. Also note that the QAM rings are unevenly spaced, with some quite close together, which makes it even more difficult to compensate for nonlinearities.

In fiber optics, APSK can also be employed in nonlinear noise scenarios and for improving tolerance against nonlinear fiber characteristics. For speeds of 400 Gbps and beyond, however, 16-QAM is preferable because of its easier implementation and better optical signal-to-noise-ratio performance due to the larger distance between the constellation points.

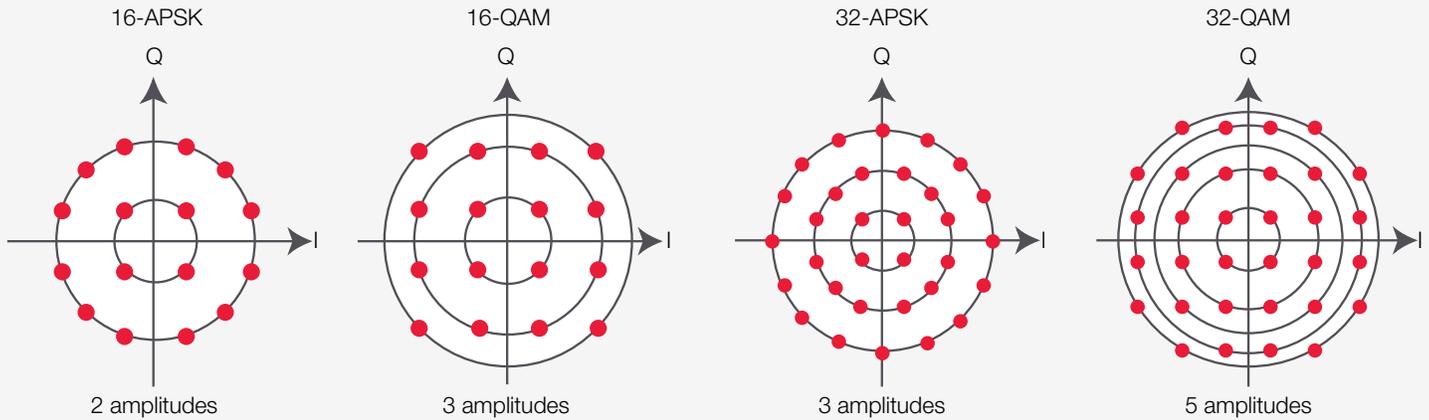


Figure 19. Constellation diagrams for APSK schemes and corresponding QAM formats

The table below provides the characteristics of the modulation formats discussed by the Optical Internet Forum (OIF) for 400 Gbps for non-polarization-multiplexed signals. If PDM is applied as well, the result is half the channel spacing or double the spectral efficiency.

| Modulation format | Coding efficiency | Symbol rate per lane/ wire | Number of carriers | Channel occupancy | Spectral efficiency (bits/s/Hz) | Maximum reach |
|-------------------|-------------------|----------------------------|--------------------|-------------------|---------------------------------|---------------|
| 64-QAM | 6 bits/ symbol | 42.7 GBd | 1 | 50 GHz | 8 | ~ 100 km |
| 16-QAM | 4 bits/ symbol | 64 GBd | 1 | 75 GHz/ 100 GHz | 5.3/4 | |
| 16-QAM | 4 bits/ symbol | 32 GBd | 2 | 75 GHz/ 100 GHz | 5.3/4 | < 1,000 km |
| 16-QAM | 4 bits/ symbol | 64 GBd | 1 | 75 GHz/ 100 GHz | 5.3/4 | |
| 8-QAM | 3 bits/ symbol | 42.7 GBd | 2 | 75 GHz/ 100 GHz | 5.3/4 | > 1,000 km |
| QPSK | 2 bits/ symbol | 64 GBd | 2 | 150 GHz | 2.7 | |
| QPSK | 2 bits/ symbol | 32 GBd | 4 | 150 GHz | 2.7 | |

Table 1. Characteristics of modulation formats at 400 Gbps as discussed by Optical Internet Forum (OIF)¹

Multicarrier implementations can now aggregate channels based on a grid granularity of 6.25 GHz and a width equal to an integer multiple of 12.5 GHz — often referred to as ‘flex-grid’. The International Telegraph Union (ITU)² has defined this flexible grid that allows scalable channels — another possibility to further increase spectral efficiency.

1. Technology Options for 400G Implementation (OIF-Tech-Options-400G-01.0)

2. ITU Recommendation G.694.1 (02/12): Spectral grids for WDM applications: DWDM frequency grid

Time Domain Pulse Shaping for Increased Spectral Efficiency

In the pursuit of using existing limited bandwidth resources most efficiently, one has to be aware that the complex modulated signal spread in time and that consecutive symbols may overlap, which is referred to as inter-symbol interference (ISI). ISI induces errors in the interpretation of the signal on the receiver side. In the frequency domain as well, care needs to be taken to avoid interference between adjacent channels. For data rates of 400 Gbps and 1 Tbps, this will become a topic of crucial importance.

This chapter elaborates on the conditions needed to prevent these effects and on the use of different filtering techniques for the purpose of bandwidth and signal containment.

Nyquist ISI criterion

The Swedish engineer Harry Nyquist explained in the 1920s that for eliminating ISI, the impulse response $h(t)$ needs to fulfill the following requirement in the time domain¹:

$$h(nT_s) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

– for all integers n . T_s is the pulse spacing of adjacent pulses.

1. H. Nyquist: Certain Topics in Telegraph Transmission Theory, Trans. AIEE, Vol. 47, pp. 617-644, Apr. 1928

In Figure 20, a signal that meets this condition—sinc (t) pulses—illustrates the impact of the criterion: The pulses overlap, but only the sampled symbol contributes to the response at the sampling instant t_k . The other symbols are zero at this point in time. In this way, the signal-degrading and error-inducing effects of ISI are avoided.

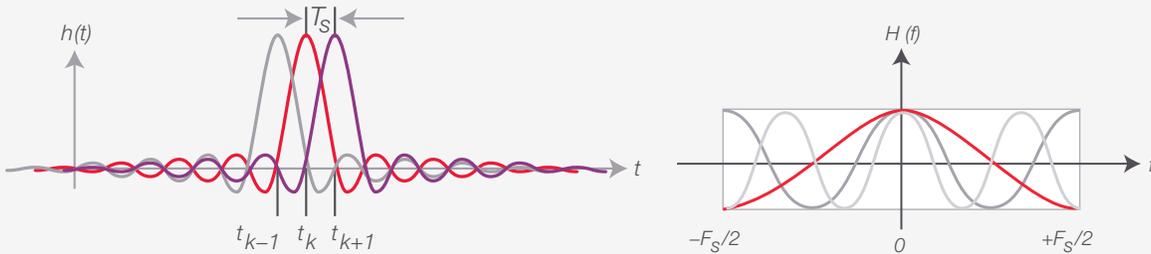


Figure 20. Orthogonal sinc (t) = $\frac{\sin(t)}{t}$ pulses meet the Nyquist ISI Criterion. ¹

The right side of Figure 20 shows the Fourier transform (FT) of the impulse response. It can be seen that the frequency response that fits into a rectangular frequency window fulfills the Nyquist ISI criterion:

$$F_s \sum_{k=-\infty}^{+\infty} H(f - k \cdot F_s) = 1 \quad \text{with} \quad F_s = \frac{1}{T_s}$$

This means that the harmonics—the components with a frequency that is an integer multiple of the so-called Nyquist frequency F_s —must add up to a constant value, in order to fit into a frequency band without ISI. The Nyquist frequency F_s is the minimum bandwidth needed to encode a signal without loss of information.

Nyquist pulse shaping using a finite impulse response filter

A sinc-signal may be perfect for preventing ISI as described above, but it is not practical because it is infinitely extended in time. Therefore, it needs to be truncated in the time domain, which is performed in practice by using a finite impulse response (FIR) filter. An FIR filter of order R responds for $R+1$ sample points and then returns to zero. Only past sample points $x[n-l]$ are considered for the convolution of the filter output $y[n]$ so that filtering can be performed in real time.

¹ Technology Options for 400G Implementation (OIF-Tech-Options-400G-01.0)

The discrete-time FIR filter output $y[n]$ in dependence of the input $x[n]$ is described by:

$$y[n] = \sum_{i=1}^R b_i x[n-i]$$

where b_i represents the filter coefficients.

To avoid aliasing, the pulse-shaping FIR filter must oversample for example by a factor of $q = 2$. In other words, there must be more than 1 sample point within T_s , so that the pulse shape can be reconstructed on the receiver side, without losing higher-frequency components.

Figure 21 shows the filter results for a sinc pulse filtered at different filter order R , always oversampled by a factor of $q = 2$. The power spectrum results from the convolution of the rectangular shaped spectrum of the sinc-pulse with the sinc-shaped spectrum of the rectangular window.

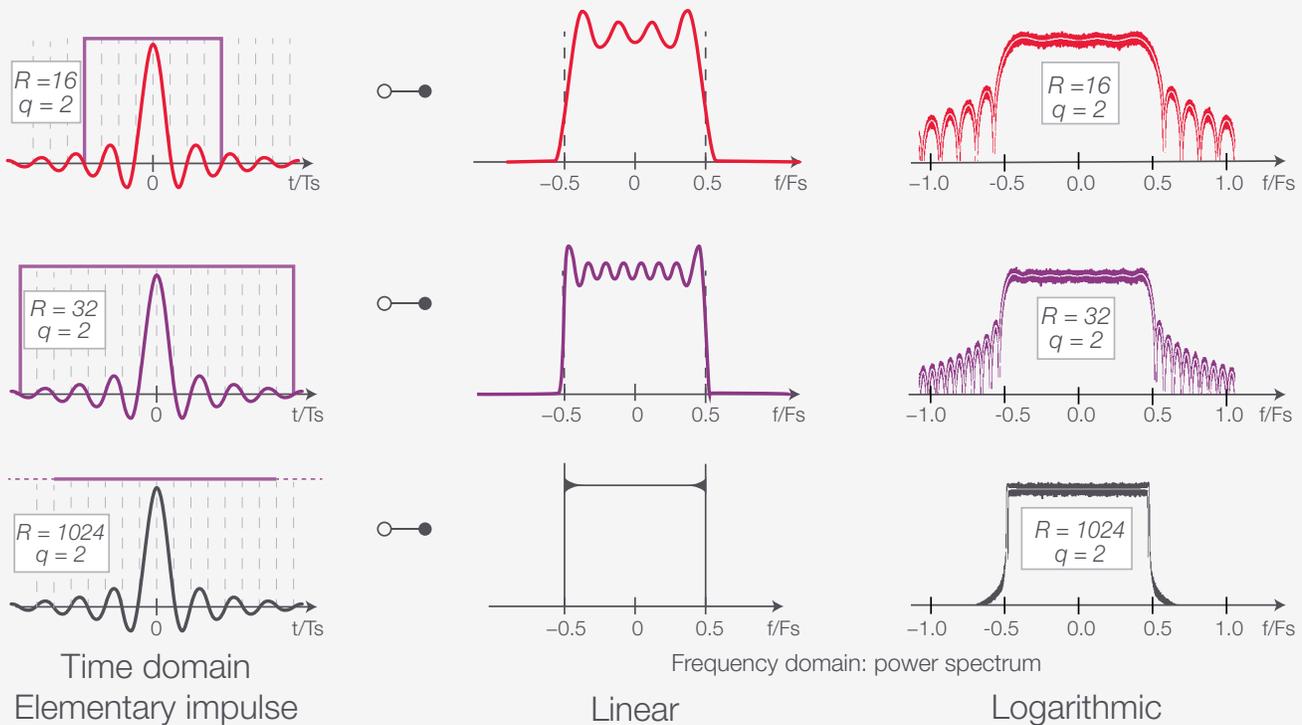


Figure 21. FIR filter of different order R used to truncate sinc-signal: time domain waveform, after Fast Fourier transformation power spectrum on linear and on logarithmic scale¹

1. R. Schmogrow M. Winter, M. Meyer, D. Hillerkuss, S. Wolf, B. Baeuerle, A. Ludwig, B. Nebendahl, S. Ben-Ezra, J. Meyer, M. Dreschmann, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold, Real-Time Nyquist Pulse Generation Beyond 100 Gbit/s and its Relation to OFDM, Optics Express, Vol. 20 (1), pp. 317 – 337, Jan. 2012

In the first line, with a filter order of 16, the signal spans $8 T_s$. In the FFT, you can see distortions caused by the limited time window. Most of the power lies within the Nyquist band ($-0.5 F_s$ to $0.5 F_s$) but a part lies outside the band. The power spectrum shows the harmonics as image spectra.

If the length of the filter is doubled ($R = 32$), the signal better fits into the bandwidth, but there is still some ringing. At $R = 1024$, the spectrum looks almost ideal; the ringing is only visible at the steep edges and the power spectrum also shows less out-of-band contribution. Unfortunately, the higher the order R of the filter, the higher the complexity of the filter design, too. Therefore, usually the lowest order R that satisfies the requirements is preferred.

The concept of raised cosine filters

For even better out-of-band suppression and a spectrum without ringing, raised cosine filters provide a reasonable alternative. Here, the impulse response is

$$h[t] = \text{sinc}\left(\frac{t}{T}\right) \frac{\cos \frac{\pi\alpha t}{T}}{1 - \frac{4\alpha^2 t^2}{T^2}}$$

dependent on the so-called roll-off factor α which can take any value from 0 to 1:

The raised cosine filters also fulfill the Nyquist ISI criterion that only the sampled symbol contributes to the signal. All other symbols are zero at the sampling points. In comparison to sinc-shaped pulses, the raised cosine signals require more bandwidth.



Figure 22 depicts the filter response for four different roll-off factors α .

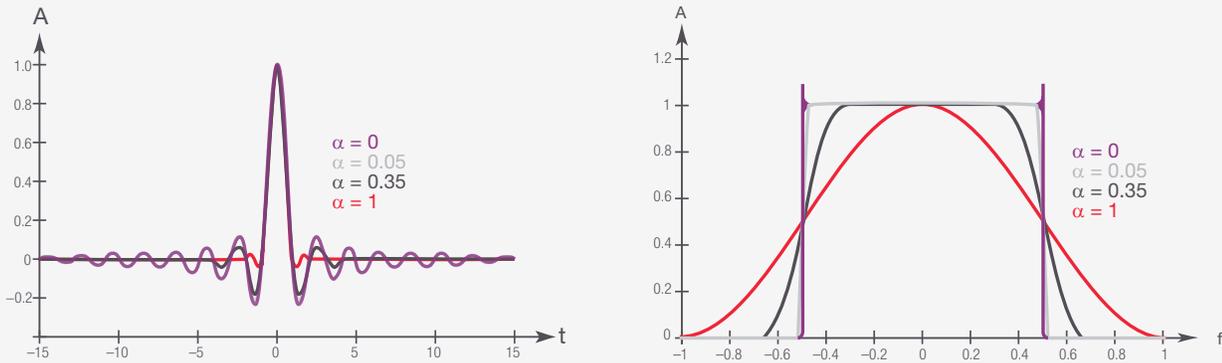


Figure 22. Raised cosine filters with different roll-off factors: normalized time and frequency domain presentation ¹

In the frequency response, it stands out that for any value of α , the curves are crossing the same point at $\pm F_s / 2$, which is half the pulse rate. As mentioned before, this is the Nyquist frequency, the minimum bandwidth needed for a data transfer without loss of information. Apart from that, for $\alpha = 1$, there is hardly any ringing, but the frequency spectrum does not fit into the bandwidth.

For $\alpha = 0$, it is vice versa: the frequency response is rectangular within the bandwidth (the overshoots at the edges are a mere mathematical effect, also known as the Gibbs Phenomenon, and do not have any practical impact). The time domain signal now shows more ringing, though.

Why should the ringing be an issue when it was first stated that at the sampling point only the sampled symbol contributes to the signal? In practice, ringing is a problem because the other symbols have zero contribution when sampling only at this ideal instant. Under real world conditions, it is almost impossible to have the receiver sample exactly at this point in time so that there will always be some ISI that induces errors in the interpretation of the signal.

Therefore, there is obviously a tradeoff between bandwidth containment and suppression of ringing in the time domain. A compromise in the choice of an adequate value of α has to be found on an individual basis for every fiber-optical application.

1. R. Schmogrow M. Winter, M. Meyer, D. Hillerkuss, S. Wolf, B. Baeuerle, A. Ludwig, B. Nebendahl, S. Ben-Ezra, J. Meyer, M. Dreschmann, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold, Real-Time Nyquist Pulse Generation Beyond 100 Gbit/s and its Relation to OFDM, Optics Express, Vol. 20 (1), pp. 317 – 337, Jan. 2012

Raised cosine filters in practice

Let's have a closer look at the influence of different roll-off factors on the most promising modulation scheme for 400 Gbps: 16-QAM. Figure 23 shows measurements not only of the frequency domain response but also the eye diagrams and the impact on the transitions between constellation points and the constellation points themselves.

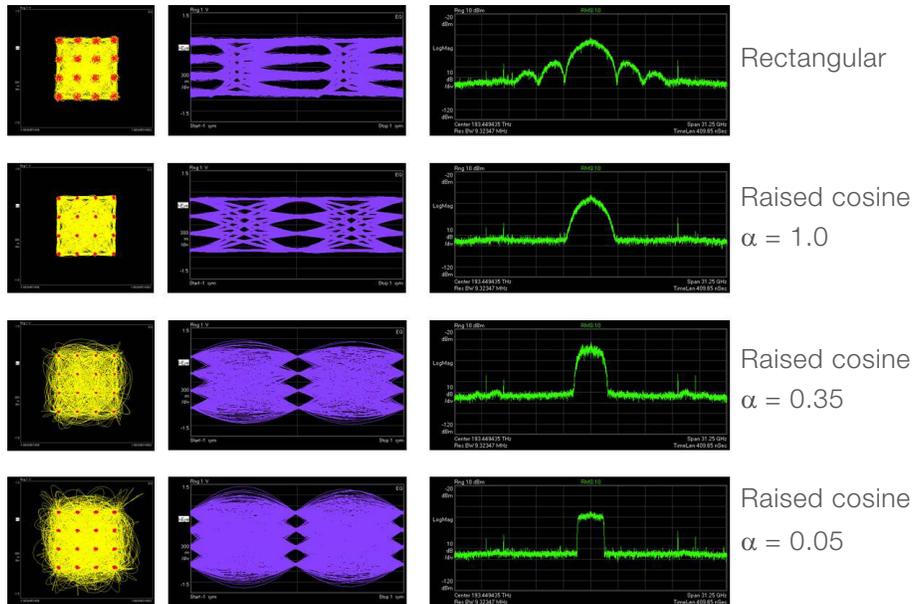


Figure 23. Raised cosine filters on 16-QAM signal in dependence of roll-off factor: constellation diagram, eye diagram and frequency spectrum; signals created with a Keysight M8190A arbitrary waveform generator

The top example shows the case of non-shaped rectangular pulses. A signal that only occupies a fixed time interval is known to have an infinitely extended frequency spectrum; in the frequency response large side lobes can be noticed. The eye diagram shows the typical behavior of a wide-band signal with open eyes. Between the constellation points, there are straight transitions.

Using a raised cosine filter with a roll-off factor $\alpha = 1$, the frequency spectrum becomes narrower; the side lobes are not visible anymore. The eye diagram shows wide-open eyes. The constellation points are smaller. This is typical for a system with reduced bandwidth. The detection bandwidth on the receiver side is, by implication, also reduced, which lowers noise.

At a roll-off factor $\alpha = 0.35$, the frequency width has further decreased, and with it the size of the constellation points. The transitions between the constellation points start to show much overshoot. This is because when reducing bandwidth, the transitions between the symbols get extended in time, which is reflected in the constellation diagram by the long, curved transitions between the points. The eyes are closing and therefore the sampling timing gets more critical.

An almost perfect rectangular spectrum is reached at $\alpha = 0.05$. The transitions between the little constellation points show large overshoot. The completely closed eyes indicate that for avoiding errors, the sampling point has to be adjusted precisely.

How much spectral efficiency can be gained?

To get an idea of the quantitative gain in spectral efficiency by pulse-shaping filters, let's compare it to the results reached by applying orthogonal frequency-division multiplexing (OFDM). Figure 24 gives a brief recap of the OFDM principles, which are similar to the Nyquist format.

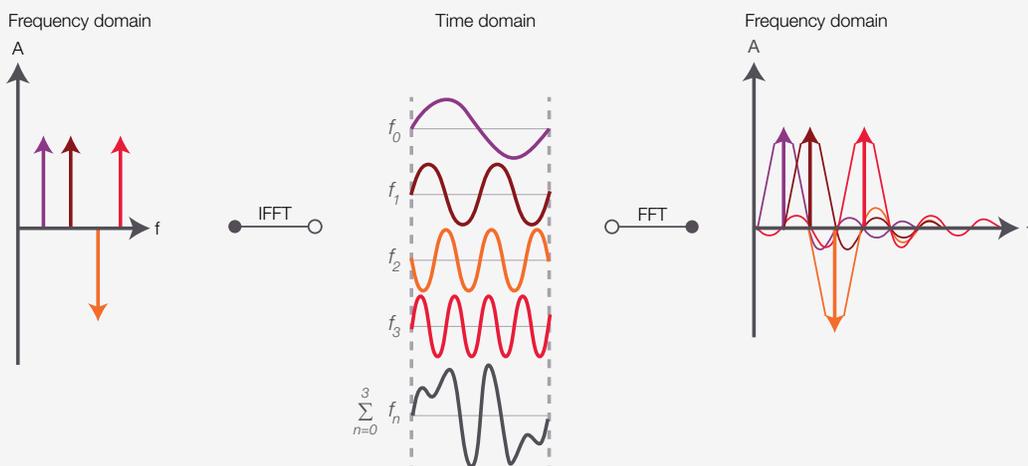


Figure 24. OFDM in the frequency and in the time domain¹

In OFDM, the frequency sub-spectra are sinc-shaped. For increased spectral efficiency, they are overlapping but because of their orthogonality — meaning that they are shifted by multiples of $\pi/2$ — they do not interfere with one another. In the time domain, a symbol is a sum of sine curves with equidistant carrier frequencies f_n in a fixed time window. In this example, One channel contains four subcarriers at four frequencies. The orange is phase shifted by π , as can be seen after inverse Fast Fourier transformation (IFFT).

1. R. Schmogrow M. Winter, M. Meyer, D. Hillerkuss, S. Wolf, B. Baeuerle, A. Ludwig, B. Nebendahl, S. Ben-Ezra, J. Meyer, M. Dreschmann, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold,.: Real-Time Nyquist Pulse Generation Beyond 100 Gbit/s and its Relation to OFDM, Optics Express, Vol. 20 (1), pp. 317 - 337, Jan. 2012

Figure 25 now shows spectral analysis on a measured 16-QAM modulated OFDM signal.

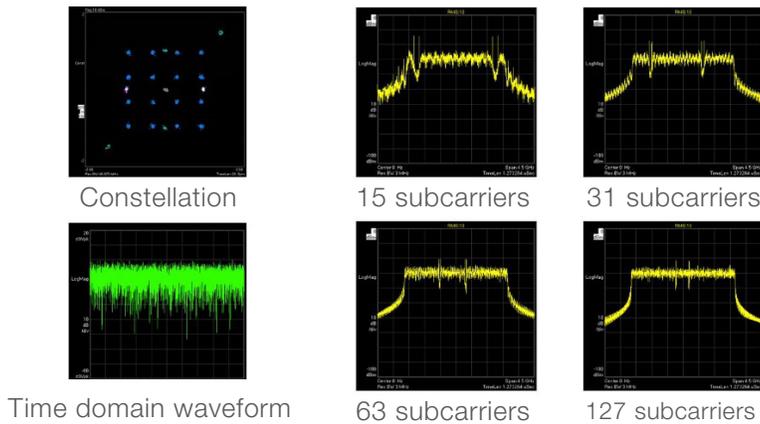


Figure 25. Impact of OFDM on a 16-QAM signal spectrum in dependence on the number of subcarriers; signals created with a Keysight M8190A arbitrary waveform generator

On the top left side, there is the constellation diagram and below the time domain waveform. With 15 subcarriers and two pilots, you can see a fairly flat frequency spectrum and a steep roll-off.

By increasing the number of subcarriers, the spectrum is flattening and the two pilots are moving towards the center. At the bottom line, it can be noticed that the spectrum is approaching rectangular shape with the growing number of subcarriers.

How does this translate now into a gain of spectral efficiency in comparison to that reached by Nyquist pulse shaping? For the sake of comparability, in Figure 26, the normalized spectral efficiency (SE) is plotted over the Nyquist filter length R (the oversampling factor q is chosen again to be 2) versus the number of OFDM subcarriers N .

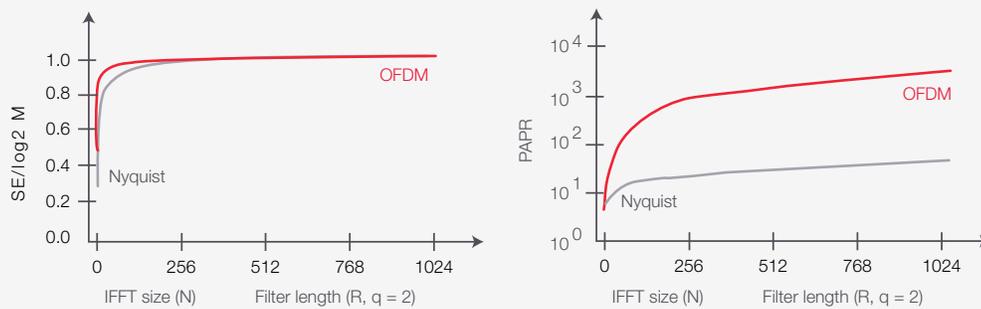


Figure 26. Effect of Nyquist pulse shaping versus OFDM on spectral efficiency and peak- to-average power ratio (PAPR)¹

1. Schmogrow M. Winter, M. Meyer, D. Hillerkuss, S. Wolf, B. Baeuerle, A. Ludwig, B. Nebendahl, S. Ben-Ezra, J. Meyer, M. Dreschmann, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold, : Real-Time Nyquist Pulse Generation Beyond 100 Gbit/s and its Relation to OFDM, Optics Express, Vol. 20 (1), pp. 317 – 337, Jan. 2012

The figure shows that both techniques provide almost the same result regarding spectral efficiency.

The comparison of the also normalized peak-to-average power ratios (*PAPR*) reveals similar behavior at differing levels. The *PAPR* of OFDM time domain waveforms is much larger. This phenomenon owes to the fact that in OFDM, the signal exhibits some peak values high above the average power level. As a consequence, OFDM circuits and test instruments need a large dynamic range to avoid distortions induced by clipping the higher levels.

Up to this point, the technical requirements for the more sophisticated data transmission concepts have been touched several times. The next chapter will investigate in detail the challenges imposed by complex modulation for the technical implementation of an optical transmitter.

An Optical Transmitter for Every Need

The previous chapters showed that by leaving behind OOK where only the amplitude of light is considered as a carrier of information, a new world of possibilities opens by also using the phase of the light wave for encoding data. It was mentioned that this goes hand in hand with technical challenges.

For transmission of OOK signals, a laser directly modulated by the electrical signal is all that is required. The result is a light signal of binary intensity. Such a low-cost and compact approach is difficult to find when the phase has to be modulated, too.

Controlling the phase with the electro-optic effect

The big relief here is that while the transmitters must become more sophisticated, there is no need to consider dispersion compensation throughout the network. Dispersion impairments can now be managed via signal processing algorithms at the receiver side, which saves a lot of money in the construction of new optical networks and line systems because no longer dispersion compensation modules are needed.

When building a phase modulator, one can benefit from the effect that the refractive index n of certain crystals such as lithium niobate depends on the strength of the local electric field. That's known as the "electro-optic effect."

How does this help for phase modulation? if n is a function of the strength of the field, then so is the speed and wavelength of the light traveling through the crystal. Thus, if a voltage is applied to the crystal, then the wavelength of the light crossing the crystal is reduced and the phase of the exiting light can be controlled by choosing the adequate voltage (see Figure 27).

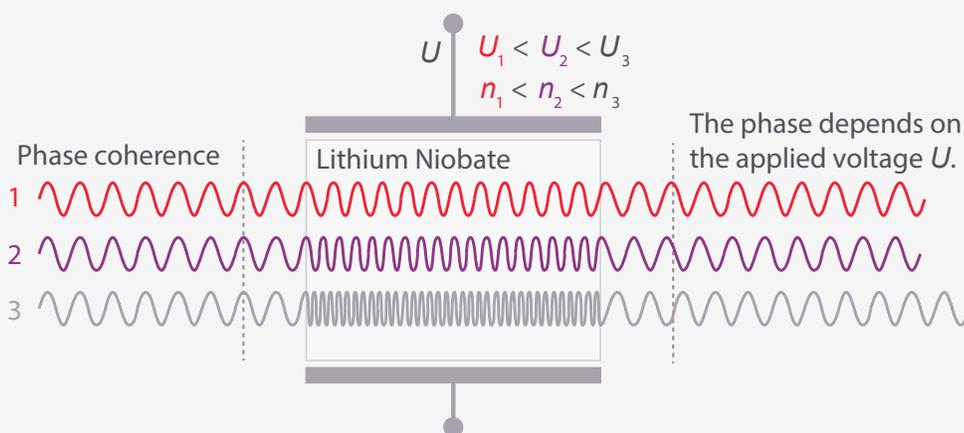


Figure 27. The higher the applied voltage U is, the slower the light travels through the crystal. This can be used to control the phase of the exiting light

In practice, this effect is used in so-called Mach-Zehnder modulators. Here, a light beam is split in two and one or both of the resulting beams travel through a phase shifting lithium niobate element such as just described. Afterwards, the two signals have a phase difference $\Delta\phi$ that depends on the voltage difference ΔU between the two optical paths. Hence, when they're recombined, ΔU also determines if they interfere constructively or destructively.

Figure 28 shows the block diagram and periodic relation between transmitted light power and ΔU . The half-wave voltage U_π is the voltage needed for a phase change of π in power transmission, meaning the voltage difference between the modulator transmitting no power and transmitting 100% of the input power.

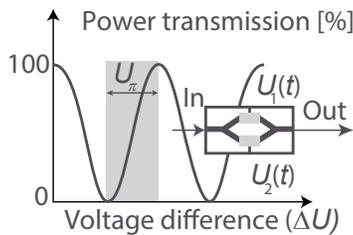


Figure 28. Periodic relation of transmitted power and voltage difference in a Mach-Zehnder modulator, with U_π being the voltage difference between zero transmission and full power transmission¹

¹ P.J. Winzer and R.J. Essiambre: "Advanced Optical Modulation Formats," IEEE Proceedings, 94(5), 2006.

The phase shifting effects of a Mach-Zehnder modulator also can be depicted in an IQ diagram. In Figure 29, you can see the example of a sine wave of constant amplitude and define relative phase $\Phi = 0$. After the split of the signal, there is only half the power on each branch. In the purple example, no voltage is applied to the modulator branches so the relative phase of the signals stays unchanged on both arms if they have equal length. The recombination results in the same sine wave with the original amplitude.

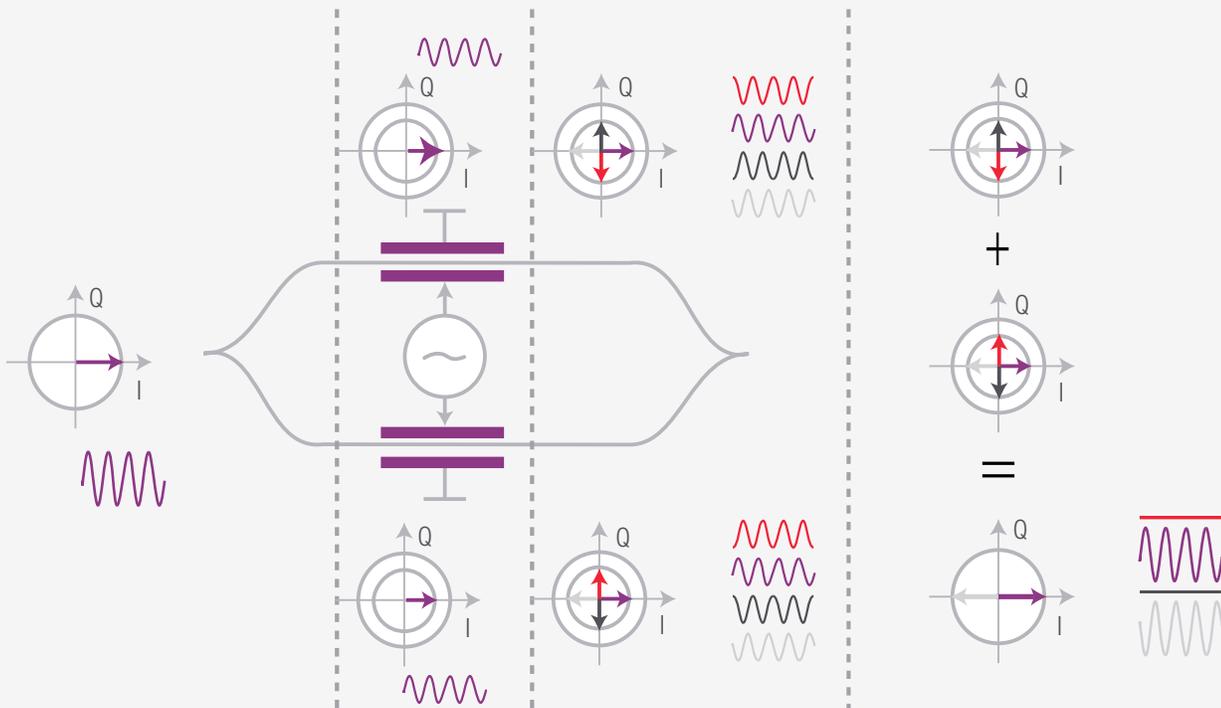


Figure 29. Examples of phase shifts in a Mach-Zehnder modulator, time domain and IQ diagram

In the red example, the signal on the lower branch experiences a phase shift of $\pi/2$ and on the upper branch a phase shift of $3\pi/2$. In the dark gray example, it's vice versa. Both examples have in common that when recombining the signals of both arms, they interfere destructively; the two vectors sum up to a zero vector. Therefore, in the red and dark gray examples, there is no signal at the exit of the modulator.

In the case of the light gray signal, the Mach-Zehnder modulator's voltages are adjusted in such a way that on both branches the phase of the signal is shifted by π . When overlaying the two signals, you get a constructive interference. The resulting wave is a sine wave of the original amplitude shifted by $\Phi = \pi$.

Mach-Zehnder modulator for transmission of a QPSK signal

How is the Mach-Zehnder modulator used in a transmitter setup using QPSK modulation? In Figure 30 the complete block diagram is given along with a recap of the principles of the QPSK modulation discussed in the chapter entitled, “Which optical modulation scheme best fits my application?”

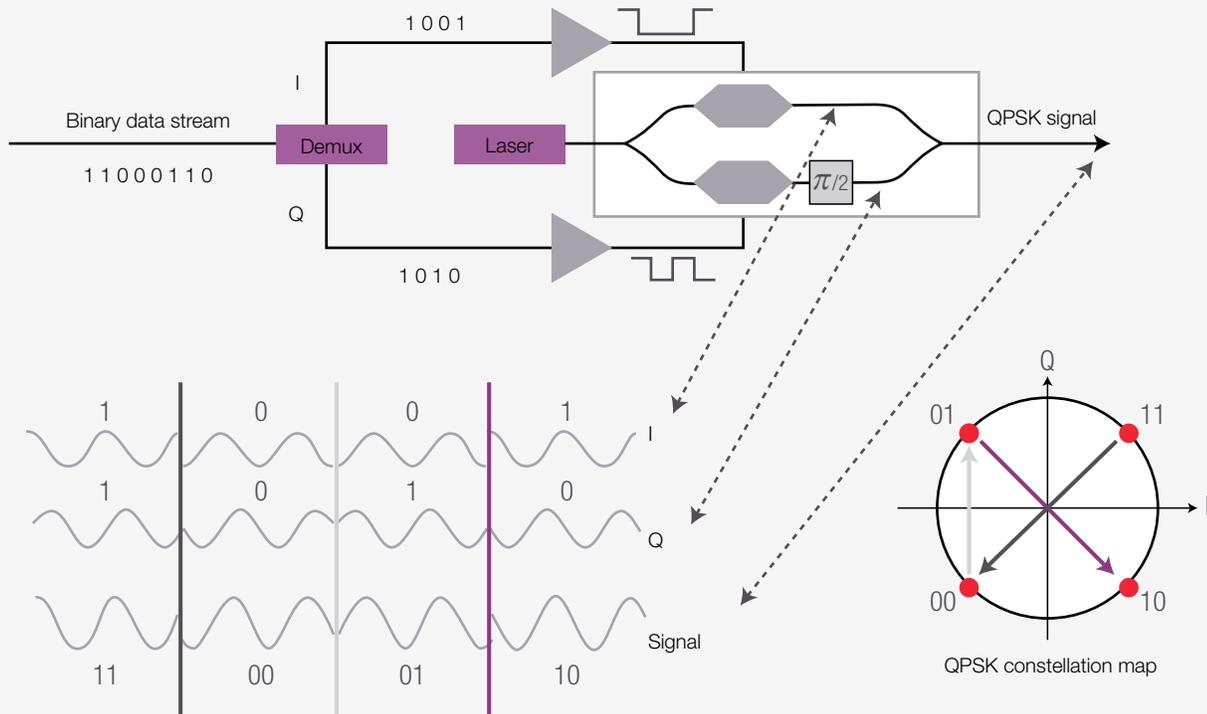


Figure 30. Transmitter setup for modulation of a QPSK signal

In QPSK modulation, the transmission rate relative to OOK is doubled by encoding 2 bits into one symbol. The four possible symbols are in the IQ diagram's four constellation points, which are all situated on the same circle. That means operating with one amplitude only. The points are separated by $\pi/2$.

In the transmitter, the electrical bit stream is split by a demultiplexer into the I and Q part of the signal. Each of the two parts directly modulates the phase of the laser signal on one arm of a Mach-Zehnder modulator. An additional Mach-Zehnder element shifts the phase of the lower branch, the Q branch, by $\pi/2$. After recombination of the two branches, the result is a QPSK signal as shown at the bottom of Figure 30.

Transmitters for more complex modulation schemes

When it comes to higher-order modulation schemes like 16-QAM, the transmitter setup must be able to provide more amplitude levels and phases, which means higher complexity. In 16-QAM, each symbol encodes 4 bits and two optical power levels are needed. To accomplish this, there are several approaches that differ in their modularity and in the level of performing the modulation in the electrical or optical domain. Figure 31 offers four implementation examples for comparison.

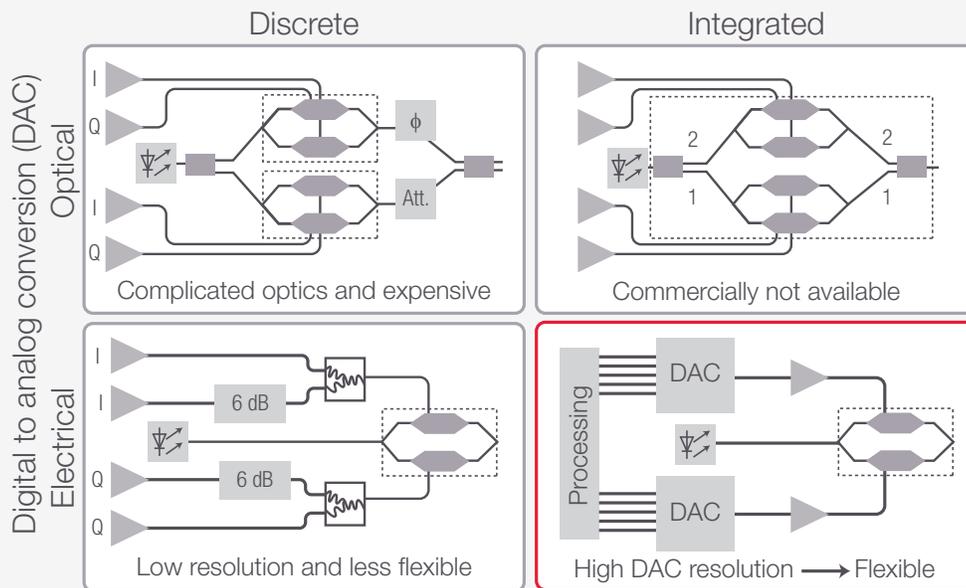


Figure 31. Transmitter for modulation formats beyond QPSK, e.g. 16-QAM. In practice the bottom right setup is commonly used

On the top left is a transmitter that consists of discrete components. The digital-to-analog conversion (DAC) is done on the optical signal. Using a BERT as an example, there are four output channels to electrically create the 4 bits of a symbol. The four voltages drive two Mach-Zehnder modulators. A laser source with a following splitter provides the two light signals that are then modulated by the Mach-Zehnder interferometers. On the lower branch, an optical attenuator follows to get the second lower light amplitude. The upper branch has another Mach-Zehnder element for shifting the phase relative to the lower branch. After recombination, the result is the 16-QAM optical signal as the interference signal. The fact that more than one Mach-Zehnder modulator is needed, is the weakness of this setup because they're expensive components. At the same time, the lithium niobate elements must have the same constant working temperature to enable precise phase control. That's difficult to guarantee.

The phase would be easier to control if the Mach-Zehnder interferometers were integrated on one optical chip as in the example in the top right of Figure 31. Here, branches 1 and 2 are each emitting a QPSK signal. The interference result of both branches is the 16-QAM signal as shown in Figure 32. The drawback to this approach though is that it's not commercially available.

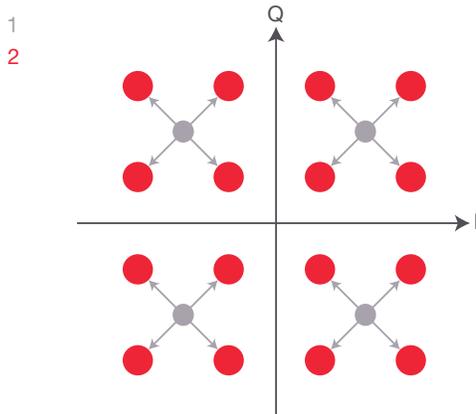


Figure 32. 16-QAM modulation in two parallel steps: on one branch, the light gray QPSK signal is modulated and convoluted with another QPSK signal on the second branch which gives the red 16-QAM constellation points

In the bottom left example of Figure 31 there are two designs where the DAC is performed in the electrical domain. A standard pattern generator can be used for creating the 4-bit sequences. The *I* part of the signal is on the upper two arms. An attenuator in one arm provides a second amplitude level. The same situation exists on the lower two arms where the *Q* part of the signal has its origin. After passing a combiner, the two-level electric signals control the optical signals of a Mach-Zehnder modulator. The disadvantage of this approach is that because of its many components the setup is very complex and therefore not flexible. It is also not possible to realize a higher-order modulation scheme like 64-QAM because the voltage resolution is not good enough for more than two amplitude levels.

The bottom right block diagram of Figure 31 shows the most convenient and flexible variant. In practice this is the typically used transmitter implementation. The signal is modulated for example with an arbitrary waveform generator, which then drives a Mach-Zehnder interferometer. With this approach it is no problem to generate more levels. Far more complex modulation schemes than 16-QAM can be put into practice with this kind of optical transmitter.

How to Detect Complex Modulated Optical Signals

In the last chapter it was shown how the use of complex optical modulation schemes affects the transmitter architecture. It is not surprising that also on the receiver side new concepts are needed.

In OOK, the signal can be detected simply by a photodiode, which converts the optical power into an electrical current I_{photo} . The photocurrent I_{photo} originating in the photodiode is directly proportional to the product of the optical signal S and its complex conjugate S^* . In the equation in Figure 33, the result only contains the amplitude A_S . I_{photo} does not provide any information on the angular frequency ω_s and the phase ϕ_s . Thus, the QPSK signal in the time domain on the right side cannot be directly mapped to the IQ diagram on the left without ambiguity. It is only possible to derive that the lower curve passing zero represents the diagonal transitions between the four constellation points, and that the middle curve represents the outer transitions. The flat signal through 1 represents the cases where the phase does not change, meaning that a symbol is followed by the same one.

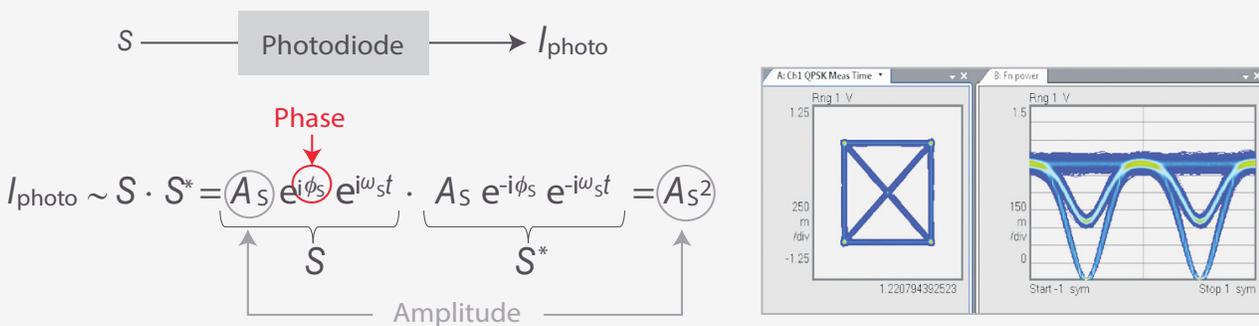


Figure 33. In direct detection, the photo current I_{photo} only provides information on the light amplitude.

For unambiguous identification of the symbol transitions, more sophisticated methods that enable detection of the complete electric field including phase information are needed.

How to detect the phase of an optical signal?

Complicating the problem is the fact that in today's optical communication systems are operated at wavelengths in the near infrared – for example, at 1550 nm, corresponding to a frequency close to 200 THz. So the changes of the electrical field in time and space are several orders of magnitude too fast to be processed with available electronics operating in the MHz to GHz range.

A local oscillator helps

The key to solve both problems lies in measuring not the absolute phase but the phase relative to a known reference signal. In Figure 34 the basic detection setup is depicted; the ideally monochromatic laser that produces the reference signal R is often referred to as the “local oscillator.”

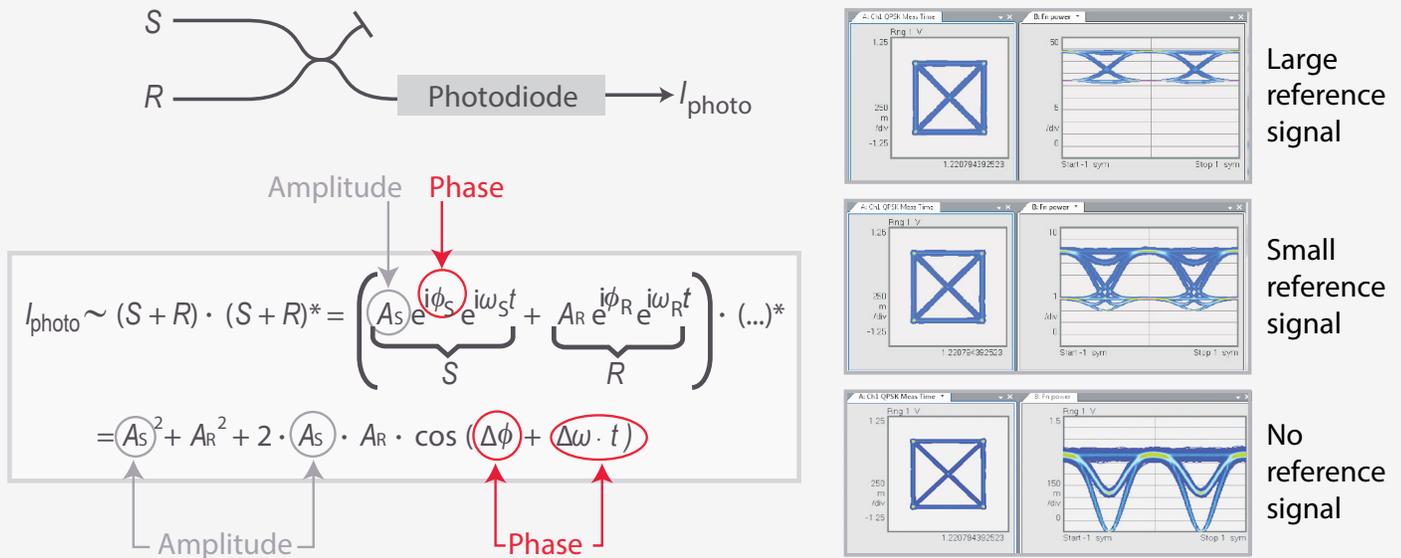


Figure 34. Mixing the signal S with a reference signal R allows measuring the phase difference. QPSK signal mixed with different reference signals. Note: We can only detect phase difference and the frequency difference shows up as linear phase change vs. time

The signal of interest S and the reference signal R are superimposed in an optical combiner and detected with a photodiode. I_{photo} is then proportional to the product of the sum of both signals $(R+S)$ and its complex conjugate $(R+S)^*$. The equation in Figure 34 reveals that the result holds the phase difference $\Delta\phi = \phi_S - \phi_R$ and the frequency difference $\Delta\omega = \omega_S - \omega_R$. From $\Delta\phi$, now the evolution of ϕ_S over time can be deduced.

The reference frequency ω_R is chosen close to ω_S so that $\Delta\omega$ is now small enough to be electronically processable.

The phase-dependent term is called the heterodyne term or beat term, because it results from mixing or “beating” the two signals.

There is also a term containing the squared amplitude that has no implications as long as only the phase is modulated and the amplitude stays constant — which is the case in QPSK modulation.

At the bottom of Figure 34 you can see the case without reference signal, discussed before, with only the A_S^2 term.

When a reference signal is added that is large in comparison to the signal itself, there is basically the beat term shifted upwards by A_R^2 . It would be advantageous to get only the beat term without this shift.

Suppressing phase-independent terms with a balanced receiver

As shown in Figure 35 all other phase-independent terms can be suppressed with a balanced receiver. Here, the signal to be detected S and the reference signal R are summed on one branch and subtracted on the second branch of a 2x2 optical combiner (which could be a fiber optical or free-space optical coupler). Each of the resulting signals is detected by one photodiode. The difference between the two photocurrents is then used. In the equation, also given in Figure 35, all other terms have cancelled out, and only the beat term remains.

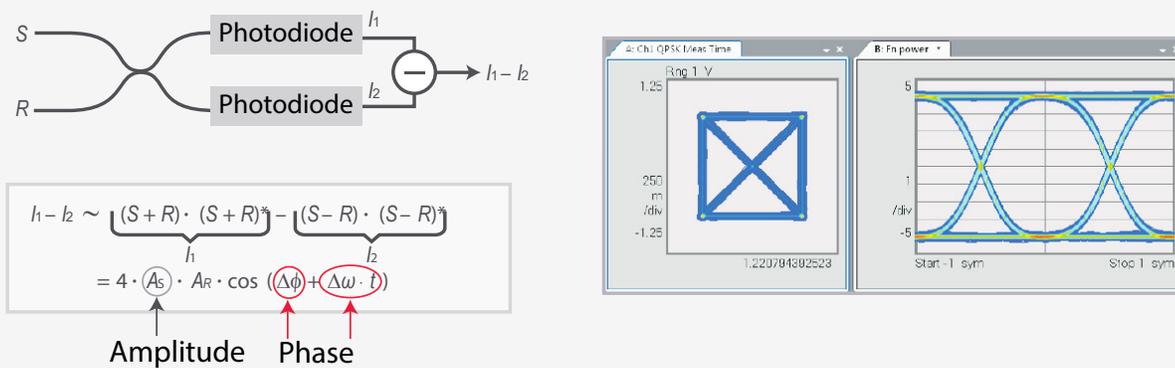


Figure 35. Using a balanced receiver, only the beat term remains with doubled intensity. Note: Balanced detection doubles the signal and removes the power of signal and reference, still we see a product of phase and amplitude information

An additional advantage of balanced detection is visible as well: the net photocurrent has doubled.

Taking the concept to the IQ plane – IQ demodulator

To recover both amplitude and phase, a coherent receiver should provide the in-phase (*I*) component and the quadrature (*Q*) component as two separate output signals. For this purpose, a second balanced detector is needed. A single local oscillator provides the reference signal for both of them, but the phase must be shifted by $\pi/2$ to obtain the *Q* part. Figure 36 gives, for the case of a QPSK signal, an idea of the whole setup, which is called an “IQ demodulator.”

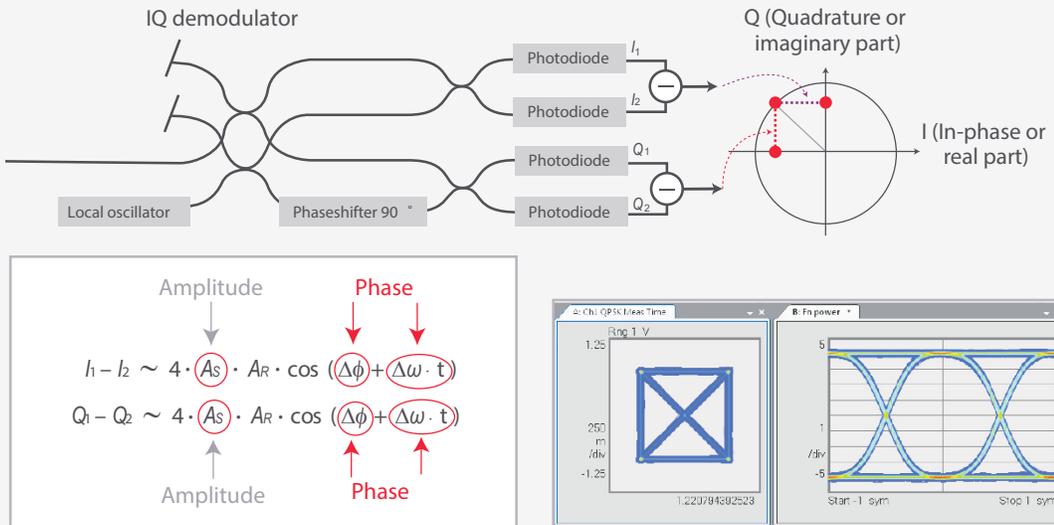


Figure 36. IQ demodulator providing two independent measurements which both contain information on amplitude and phase. Note: Adding a second balanced detector allows us to have two independent measurements that contain signal amplitude and phase which we now can recover

This setup only works for coherent signals that are not polarization-division multiplexed. In addition, the signal only mixes with the component of the local oscillator signal with the same state of polarization at the detector.

Extending the concept to dual polarization

For dual polarization, the demodulator concept needs further development. The basic principle stays the same: After a polarization splitter, there are now two IQ demodulators, one for the x-polarization and the other one for the y-polarization. Only one local oscillator provides the reference signals for all branches.

The block diagram is given in Figure 37. There are four output signals to resolve I- and Q-coordinates, one for each polarization direction respectively. In the equations, the upper indices h and v reflect the horizontal and vertical polarization state of the signal with respect to the polarization reference frame of the receiver. This polarization diversity architecture also assures that all of the signal is mixed with the local oscillator, regardless of the input state of polarization. Therefore, it is commonly used, even if the signal does not use dual polarization.

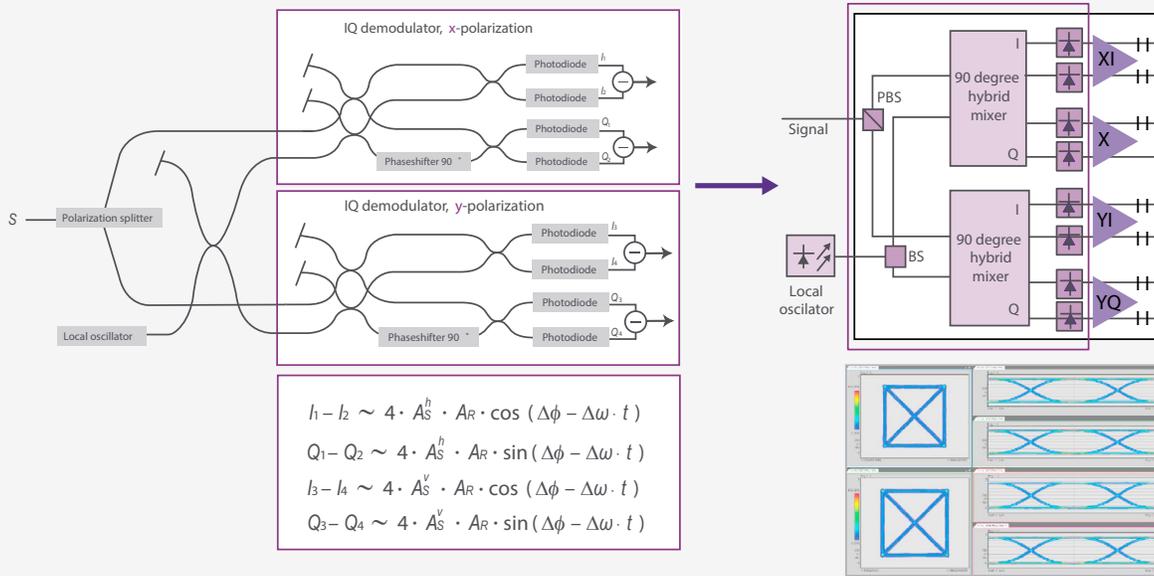


Figure 37. IQ demodulator for polarization resolved measurements¹

¹ Block diagram from "OIF Implementation Agreement for Integrated Dual Polarization Intradyn Coherent Receivers." IA # OIF-DPC-RX-01.2, Nov. 14, 2013

So far, receivers with a local oscillator of a frequency ω_R that is different from the signal's frequency ω_S have been investigated. These are called heterodyne receivers.

In homodyne receivers, the local oscillator has the same frequency as the carrier signal itself. The advantage: the above terms are not frequency dependent anymore.

Figure 38 quantifies the required electrical bandwidth for both homodyne and heterodyne receivers. For homodyne detection, where the local oscillator has the same frequency as the signal itself, half the signal's optical bandwidth is needed. For a heterodyne receiver the needed electrical bandwidth increases with the frequency offset between local oscillator and signal.

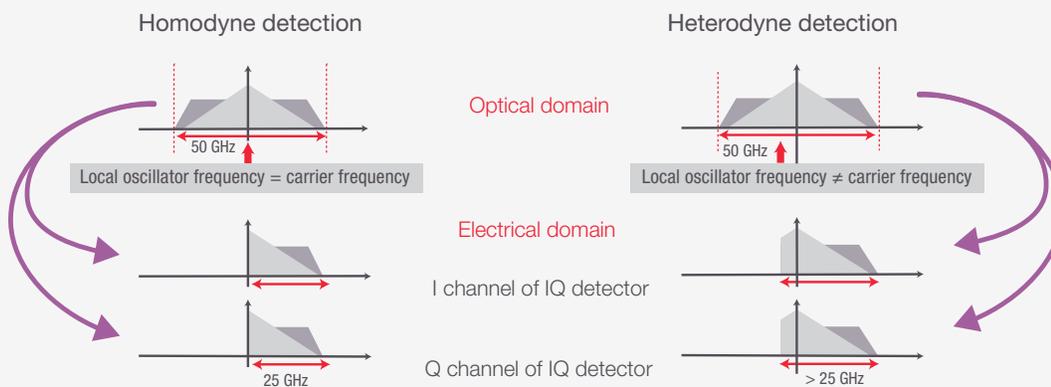


Figure 38. The required electrical bandwidth for coherent detection depends on the frequency offset between signal and its reference

Use a delayed copy of the signal as a reference – delay line interferometers

It seems that a local oscillator is indispensable for recovering the phase information. An alternative solution could be to overlay the signal with a copy of itself. This way, it is possible to have a reference signal where $\omega_R = \omega_S$.

One could think that this effort is not very promising because it may not be clear how to gain additional information on the phase that way. But this self-homodyne approach is useful because it is of interest to detect the phase change over time. So, if the signal is split in two and the signal is overlaid with the delayed copy as a reference signal, information on the phase changes can be obtained.

The advantage of this measurement method is that it is not subject to inaccuracies due to slow (in comparison to the symbol rate) frequency fluctuations of an external local oscillator and of the carrier laser itself.

This kind of receiver setup is known as a delay-line interferometer. Figure 39 shows a balanced delay-line interferometer with the signal $S(t)$ and the signal $S(t+T)$ delayed by T .

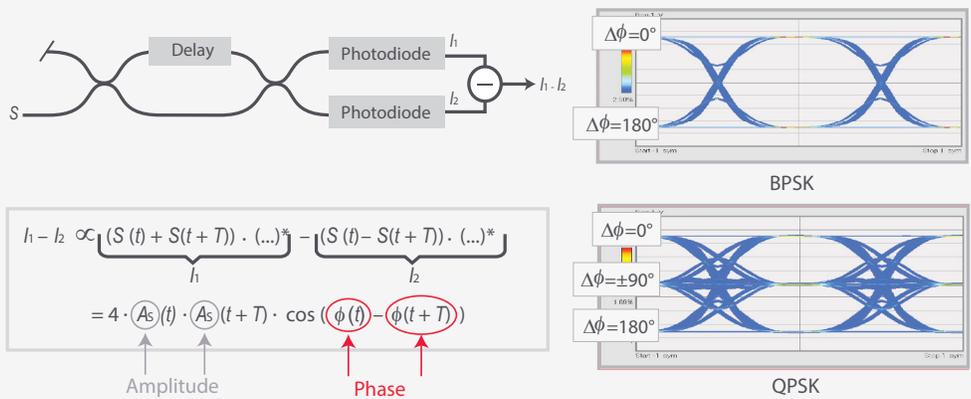


Figure 39. Balanced delay line interferometer

The equation here shows that the result is dependent on the cosine of the phase difference between the original signal and its delayed copy. Due to the periodicity of this function, only phase differences between 0 and π can be uniquely identified and only for delays T that are approximately an integer multiple of the carrier period $2\pi/\omega_s$. This is sufficient for BPSK but for phase recovery of QPSK and higher-order modulation schemes, another delay line interferometer phase shifted by $\pi/2$ relative to the other delay-line interferometer has to be added to cover the full phase range from 0 to 2π .

Figure 40 shows the setup with an additional delay-line interferometer for receiving the two independent I and Q components. $Q_1 - Q_2$ is measured additionally, while $I_1 - I_2$ stays unchanged.

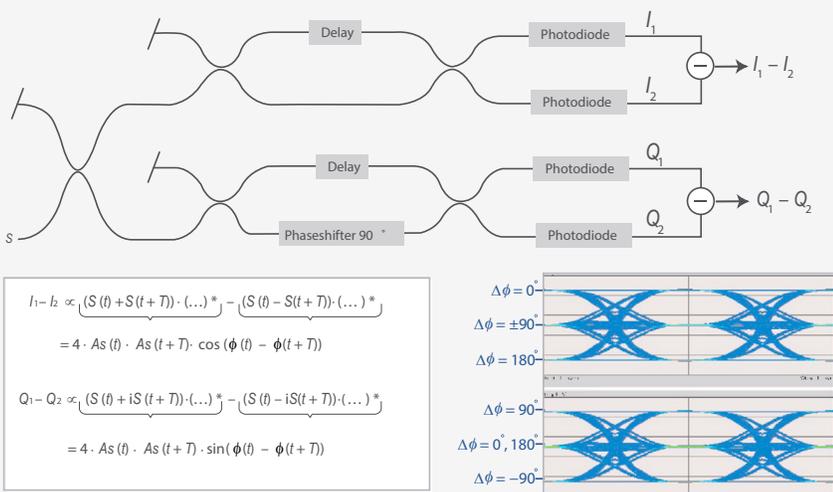


Figure 40. Extended delay line interferometer for QPSK and higher order modulation formats

Analogous to a heterodyne receiver, the delay-line interferometer can also be expanded for polarization-sensitive measurements.

With a delay-line interferometer, there is no need to have an external local oscillator. As such, oscillator-introduced phase noise is avoided and less signal processing is required. However, this approach does have disadvantages that still might lead us to prefer a heterodyne receiver.

First, to measure phase changes over time with a delay-line interferometer without clock data recovery (CDR), the delay and the sampling period need to be considerably smaller than the symbol period. Today's symbol rates have reached a level where this can be hard to accomplish. In addition, for low-power signals the measurement sensitivity is reduced because the reference signal is also of low power and suffers from noise accumulated on the transmission link. For implementation with a sampling technique, the measurement time increases and a trigger is required. Bottom line, a homodyne receiver is not very flexible.

Until now, exclusively time-domain detection techniques have been investigated. Alternatively, the frequency spectrum could be detected and from it via Fourier transformation the time-domain signal deduced.

Frequency domain detection

To recover a complex modulated optical signal from its spectrum, we have to measure the complex spectrum, meaning with amplitude and phase information.

This can be performed with a complex spectrum analyzer that separates the different optical frequency components with a dispersive element. All the frequency bands can either be detected simultaneously with multiple detectors or sequentially with a scanning narrow-band optical filter and a single detector.

For recovering phase and amplitude, again a local oscillator is employed as a reference signal. For recovering both components, a source emitting at two optical frequencies is needed.

Figure 41 shows the complete setup that is needed to measure the polarization-resolved complex spectrum.

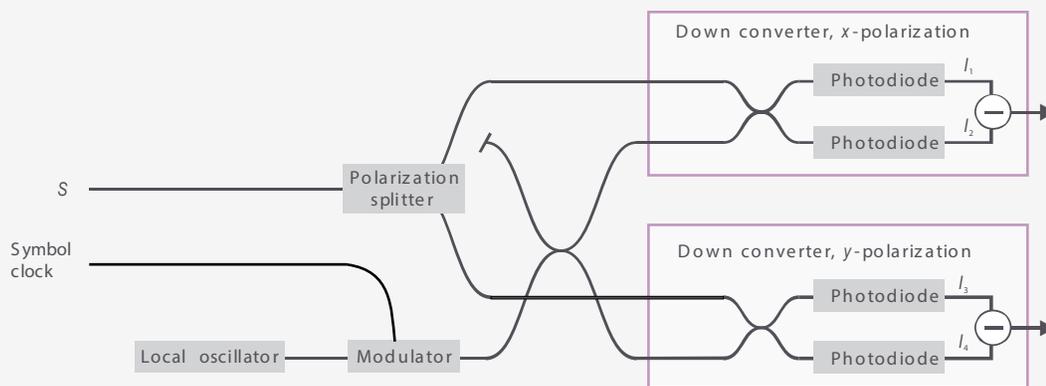


Figure 41. Setup for polarization resolved coherent frequency detection

The big plus of frequency-domain detection is its virtually unlimited bandwidth, meaning also unlimited time resolution. The bandwidth depends on the sweep range of the local oscillator so that bandwidths in the THz range can be reached with today's tunable external cavity lasers. The other big advantage is that there is no need of a high-speed receiver.

On the other hand, there are also major drawbacks.

For example, it is only applicable to periodic signals because these result in the required discrete spectral peaks. Additionally, you now need a symbol or pattern clock. The precision of the recovered time-domain signal directly depends on the spectral resolution that determines the number of sidebands that can be resolved. The spectral resolution that can be reached today limits the pattern length to a few tens of symbols.

These factors and the fact that this method does not give results in real time, make frequency-domain detection inapplicable for network receivers. In fact, this would imply long measurement times and fairly complex measurement setup and signal processing.

Finally, in frequency detection, all non-periodic effects are averaged out. This is also true for polarization-mode dispersion (PMD), which therefore cannot be compensated.

Preferences?

Self-homodyne setups need little signal processing and are the least sensitive to phase noise. Still, they are not very flexible, work only close to the design symbol rate, and are less sensitive than heterodyne implementations.

The heterodyne time domain detection methods offer the highest flexibility. Unlike frequency-domain detection, they can be used for real-time detection. They therefore are applicable to live signals in data networks. Equivalent-time sampling only works for repetitive signals of a limited length. With real-time sampling, the complete signal can be reconstructed in all domains and without limitations regarding the modulation format. There are also no limitations regarding signal length in heterodyne time-domain detection. PMD and CD can be compensated for during signal processing. In this case, only the signal processing is the throughput limiting factor.

At the same time, it has to be taken into account that four-channel high-speed equipment is needed for this approach, such as a high-performance real-time digitizer with very low jitter and noise and a high effective number of bits (ENOB) over the whole frequency range.



Coherent Optical Receivers – the Complete Answer

The last chapter showed that the most flexible detection setups are heterodyne time domain detectors, which are applicable for test and live signals and operate independent of the modulation format.

In Figure 42, such an IQ detector is positioned on the left side. The diagram makes clear that still some steps need to be taken before the incoming bits encoded in the symbols can be identified and further analyzed. The receiver architecture shown here is recommended by the Optical Internetworking Forum (OIF) and enables extraction of all information in the signal. Receiver architecture is examined in detail.

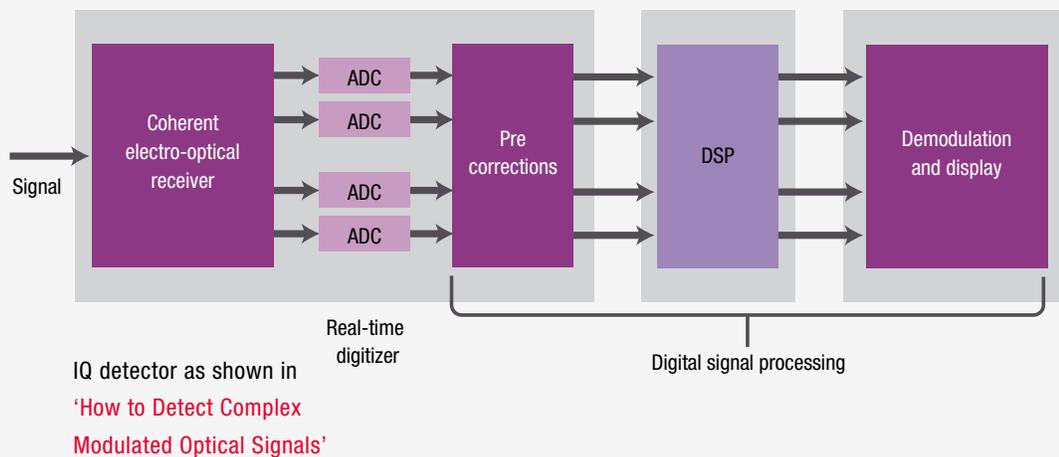


Figure 42. Receiver architecture recommended by OIF for implementation of integrated intradyne coherent receivers ¹

1. OIF Implementation Agreement for Integrated Dual Polarization Intradyne Coherent Receivers, IA # OIF-DPC-RX-01.2, Nov. 14 2013

Relieving many impairments

After analog-to-digital conversion, DSP follows as an integral part of any coherent optical receiver. The use of DSP makes the big and attractive difference versus conventional on/off keying, where the signal-distorting effects of CD and PMD exist. DSP supports algorithmic compensation for CD, PMD, and other impairments, since coherent detection provides the complete optical field information. That means complex optical modulation relieves from the need for PMD compensators or dispersion compensating fibers as well as the increase in loss and latency induced by these modules.

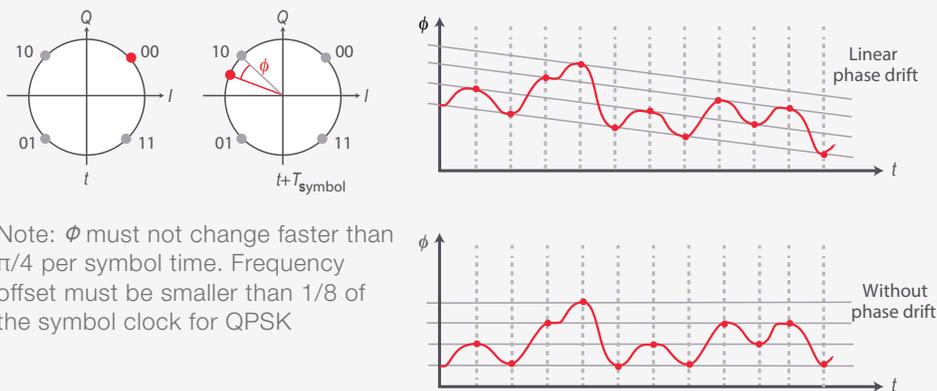
An upstream pre-corrective step removes receiver imperfections. These imperfections can be channel imbalances between the four electrical channels, IQ phase-angle errors of the IQ mixer, timing skew between the four ADC channels, and a differential imbalance of the nominally balanced receiver. For removal of these impairments, the component is typically characterized over wavelength during the instrument calibration.

Apart from these imperfections introduced by the receiver, DSP has to compensate for the degradation the signal experiences on the optical path between transmitter and receiver. These are CD and PMD, polarization-dependent loss (PDL), polarization rotation or polarization state transformation (PST), and phase noise.

For estimating the impact of phase noise, the carrier phase can be tracked over time. Nevertheless, this step is not compulsory in a coherent receiver setup.

Carrier phase recovery

By introducing an LO, a way to track the signal's phase changes over time relative to the LO's phase was found. But the fact that the LO in a heterodyne receiver scenario has a different frequency than the signal causes a linear phase drift over time. That's understood when recalling that in a heterodyne receiver I_{Photo} is proportional to $\cos(\Delta\phi + \Delta\omega t)$ ([How to Detect Complex Modulated Optical Signals](#)). Figure 43 visualizes this “rotating” constellation for QPSK modulation.



Note: ϕ must not change faster than $\pi/4$ per symbol time. Frequency offset must be smaller than $1/8$ of the symbol clock for QPSK

Figure 43. The frequency difference between the transmitter laser and the local oscillator leads to a 'rotating' constellation

For preventing ambiguities, the phase must not change faster than $\pi/4$ per symbol time, which is half the phase difference between two neighboring symbols. That in turn means the frequency offset between the LO and signal needs to be smaller than $1/8$ of the symbol clock for QPSK.

To be able to track the phase, the signal must be sampled at times with predictable phase values, for example, at the symbol times. For the case of a bandwidth-limited signal, the sampling rate of the phase is smaller than the symbol rate. In Figure 44 the dark gray line shows that the phase may not be correctly recovered.

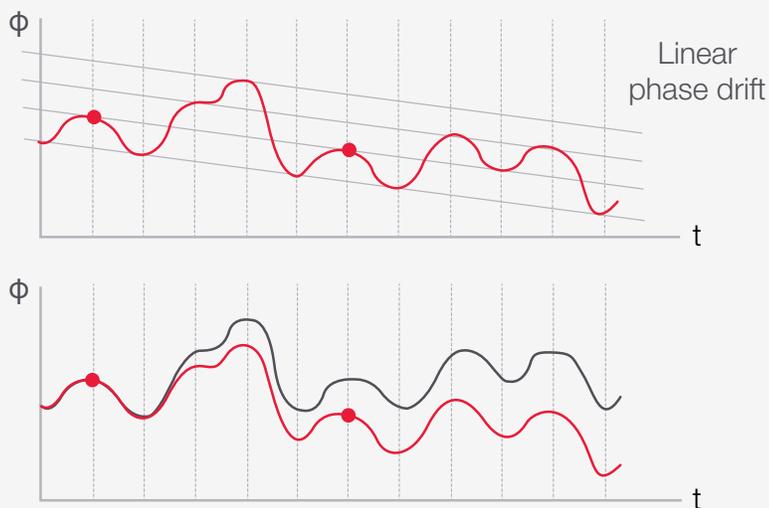


Figure 44. In a real transmission system, it is often not possible to recover the phase as the sampling rates are too little given the level of phase noise and offset

Under these circumstances, the carrier phase noise and offset must be within very tight limits to allow recovery of the phase. In a real transmission system, that's usually not the case since these tighter specifications aren't required in real line cards that use real-time acquisition.

Figure 45 shows the influence of carrier bandwidth on phase recovery for a DFB laser. On the left is an example of a high tracking bandwidth. The constellation points in the IQ plot are artificially narrow because in this case phase tracking reduces the angular width of the symbols. Lower bandwidth results in the more realistic looking round symbols. With even lower bandwidth, in the carrier phase a limit is reached where it's not possible to track the phase anymore. The angular spread of the symbols is clearly affected by phase noise that could not be removed.

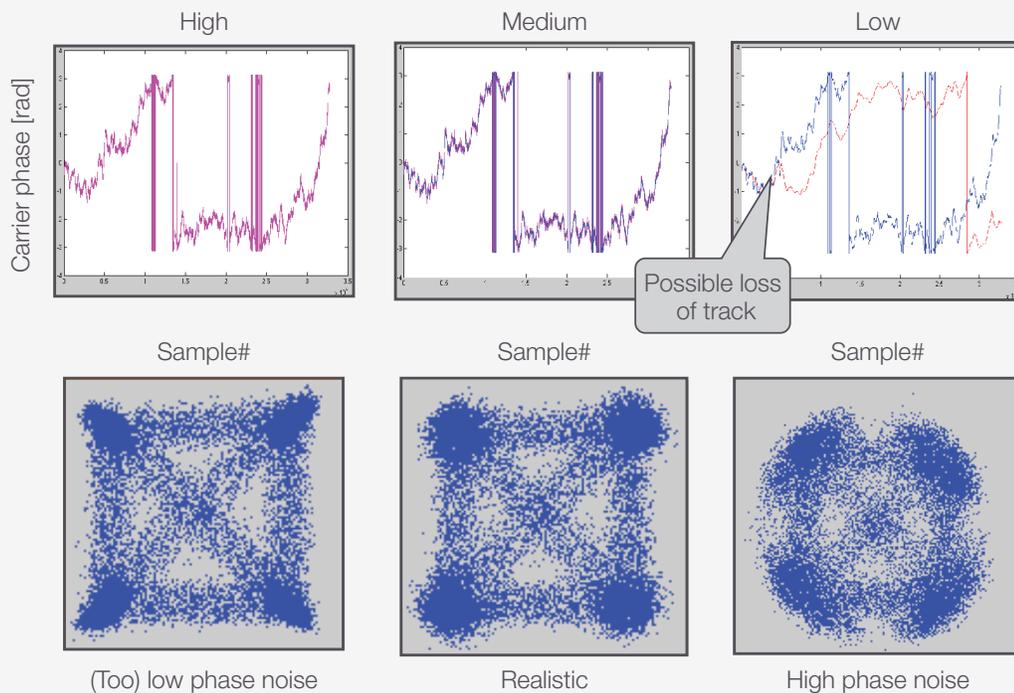


Figure 45. Examples of carrier phase tracking of DFB laser in dependence of tracking bandwidth

Finding the Jones-Matrix for recovering the original polarization state

For providing two independent baseband signals (for x and y polarization) to the digital demodulator, polarization demultiplexing is a fundamental step in DSP. In this step, compensation for PMD and PDL have to be done. It also needs to be considered that in single mode fibers (SMF) the polarization state is not preserved during propagation.

The polarization directions evolve along the signal's way through the fiber (Figure 46) so that the state of polarization (SOP) at the end is not simply related to the orientation of the receiver. Therefore, with the polarization beam splitter in the receiver, you do not get the two independent signals but a linear combination of the two polarization tributaries. Polarization maintaining fibers preserve the SOP but they are not deployed in data transmission due to their higher attenuation and price.

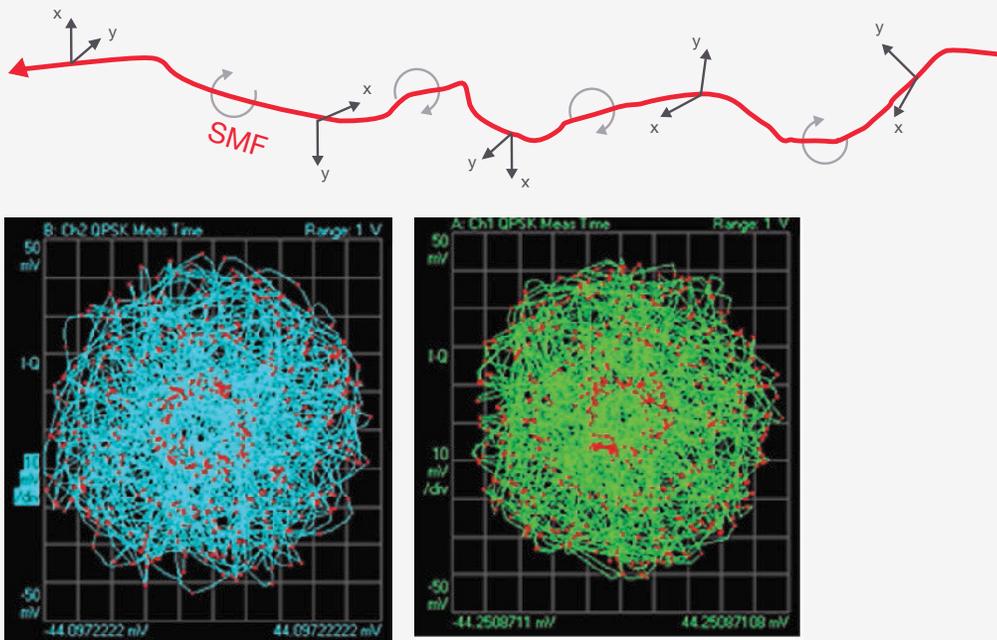


Figure 46. Single mode fibers change the state of polarization of the transmitted light. Therefore, the polarization splitter on the receiver side does not provide 2 independent signals, but a linear combination, here example of dual polarization QPSK constellation plots before polarization demultiplexing

All the degrading effects that happen to the fully polarized light in the fiber channel can be mathematically described with a so-called Jones matrix. The sent signal S is multiplied with the Jones matrix, resulting in the received signal R . For an ideal channel without impairments, the Jones matrix is an identity matrix; the received signal is the same as the originally emitted signal (see Figure 47). In its most general form, the Jones matrix is a complex 2×2 matrix with eight independent real parameters.

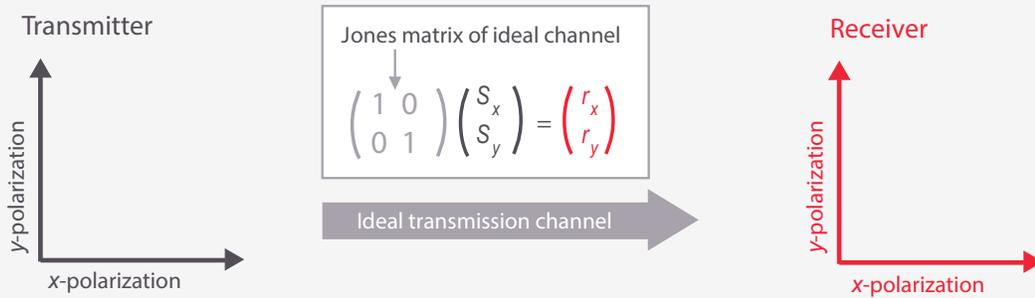


Figure 47. Jones matrix of ideal channel

So basically the Jones matrix has to be determined to deduce the original signal from the measured received signal. That's hard to do since there is usually little to no information on the quantities of the impairing effects that the signal has experienced in the channel.

Therefore, so-called blind algorithms are often employed to approximate the original signal. These are estimation techniques that don't require knowledge of the original signal (except for the modulation format). Here, a series of equalizer filters (see Figure 48), applied to the received signal, represents the inverse of the Jones matrix. Each filter element models a signal degrading effect. The algorithm iteratively searches the set of filter variables (α , β , k ...), which leads to convergence, meaning that the measured symbols map with minimal error to the symbols calculated by the algorithm.

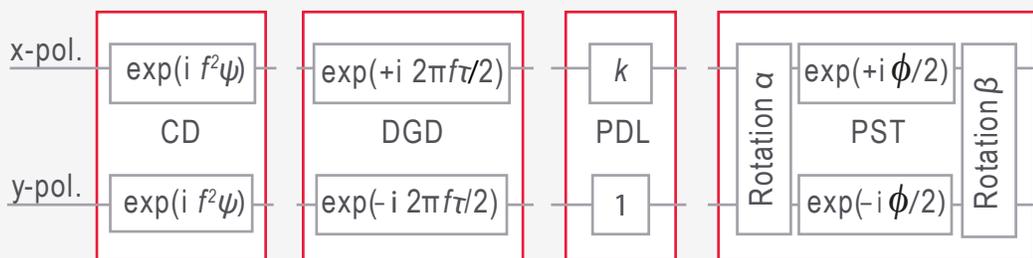


Figure 48. Model for equalizer filters compensating chromatic dispersion (CD), differential group delay (DGD), polarization dependent loss (PDL) and polarization state transformation (PST).¹

1. J. C. Geyer, F. N. Hauske, C. R. S. Fludger, T. Duthel, C. Schulien, M. Kushnerov, K. Piyawanno, D. van den Borne, E.-D. Schmidt, B. Spinnler, H. de Waardt, B. Lankl, and B. Schmauss: Channel Parameter Estimation for Polarization Diverse Coherent Receivers, IEEE Photonics Technology Letters, Vol. 20, No. 10, May 15, 2008

A drawback to this method is that it may recover the same polarization channel twice. This problem is known as a singularity of the algorithm. It is also a very complex method because each symbol has to be treated individually to calculate the next iteration step.

Easier in Stokes space

The estimation gets easier in Stokes space, where polarization demultiplexing is truly a blind procedure because neither demodulation nor knowledge of the used modulation format or carrier frequency is required. Additionally, in Stokes space there is no singularity problem.

The Stokes space helps to visualize polarization conditions of optical signals and is therefore also a good tool for observation of polarization changes along an optical channel. Any polarization state of fully polarized light can be depicted by a point in this three-dimensional space that lies on the surface of a sphere — the so-called Poincaré sphere that has its center in the origin of the coordinate system. The radius of the sphere corresponds to the amplitude of the light. Circular polarization can be found along the S3 axis. Along the equator in the plane spanned by the S1 and S2 axes, there is linear polarization, and intermediate positions represent elliptical polarization. In Figure 49, you can see in green where some of the individual polarization states are situated on this globe.

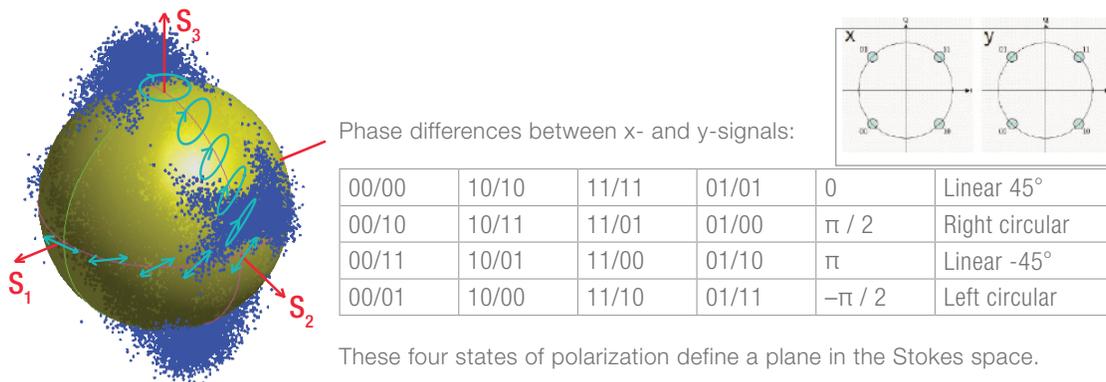


Figure 49. Poincaré sphere in Stokes space of a polarization division multiplexed (PDM) QPSK signal

Also in Figure 49, a measured x- and y-polarized QPSK signal is shown. There are four possible phase differences between the two signals at the sampling points. The combination of these x and y signals with these four phase differences gives the measured blue dot clouds in Stokes space. (With a QPSK signal in only one polarization direction, one would only get one accumulation on the S1 axis.)

The transitions between the four states define a lens-like object in Stokes space (see Figure 50). It can be shown that polarization multiplexed signals of any format always define such a lens.

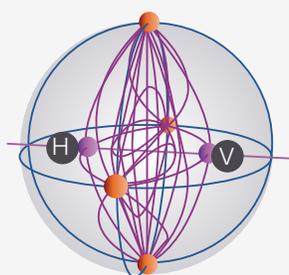


Figure 50. Transitions between the symbols of a PDM QPSK-signal, in the left hemisphere the transitions of the x-polarized (H: horizontal) and in the right hemisphere of the y-polarized signal (V: vertical)

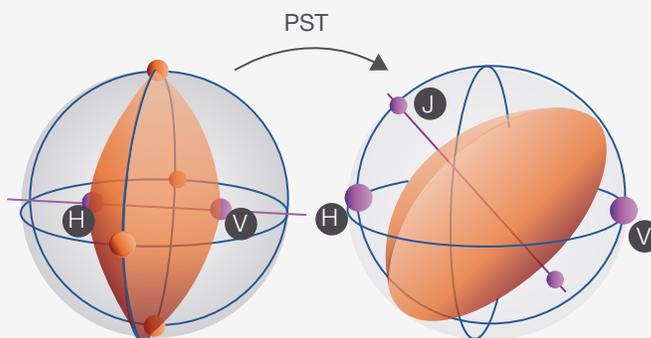
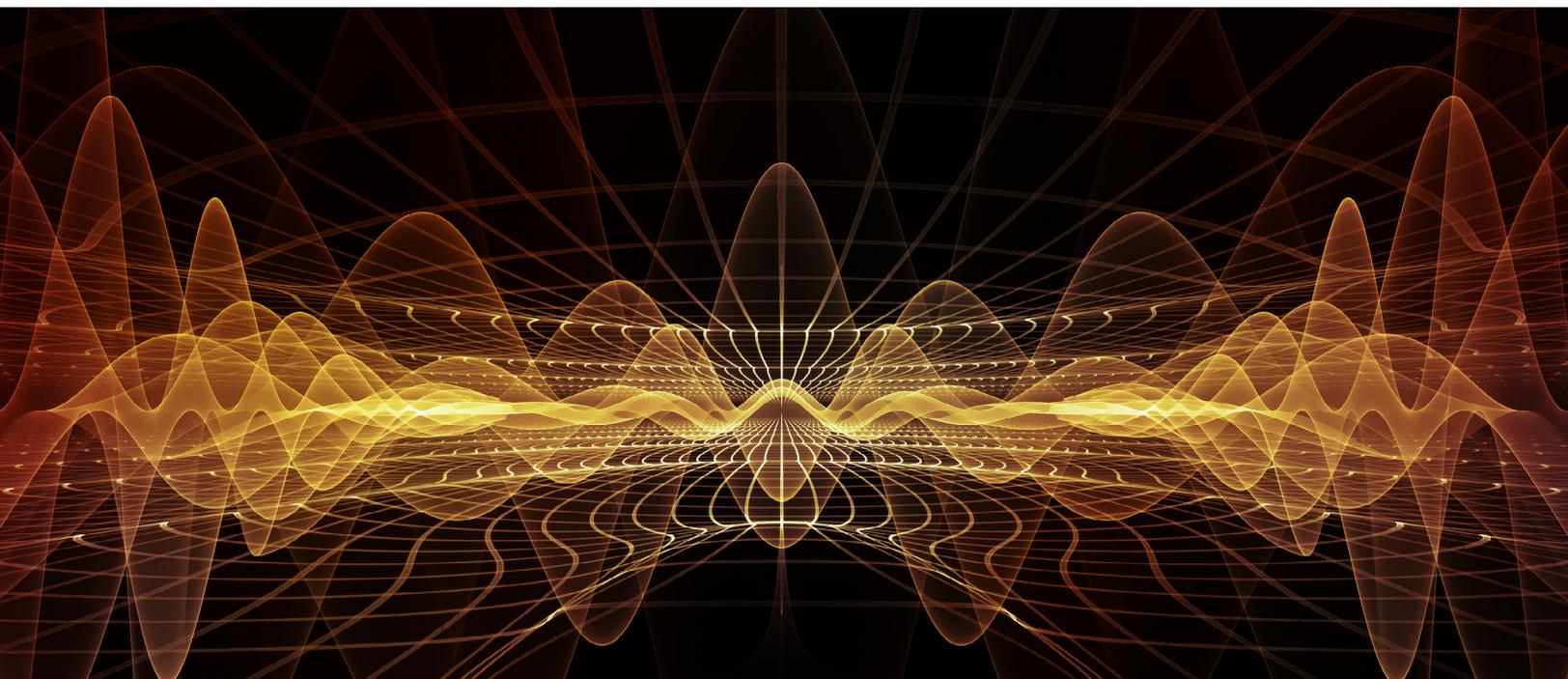


Figure 51. PST causes a rotation of the lens in the Stokes space. The normal of the lens defines the Jones matrix — here for the example of a PDM QPSK signal

When facing PST along the optical path of a single mode fiber, the lens is being rotated in the Stokes space (see Figure 51). From the rotation, the Jones matrix, which is the normal of the lens-like object, can be deduced.



How are other signal-degrading effects represented in Stokes space? With PDL, the lens is deformed and shifted. Nevertheless, that doesn't cause a problem for recovering the Jones matrix. The deformation allows quantifying the amount of PDL. CD is polarization-independent and doesn't impede polarization demultiplexing. In this case, constellation plots are the preferable tool for a quantitative investigation.

Determination of the symbols

After successful DSP, the received symbols can finally be determined. In QPSK, the decision criterion is the I and Q value of the measured point in the constellation diagram (see Figure 52), e.g., every point with positive I and Q value will be interpreted as "11." In more advanced formats, it's no longer possible to simply take the I and Q values as decision levels. The points are assigned to the closest symbol. From the diffuse clouds on the right side of Figure 52, one can tell that even with coherent detection bit errors occur. How can these be quantified?

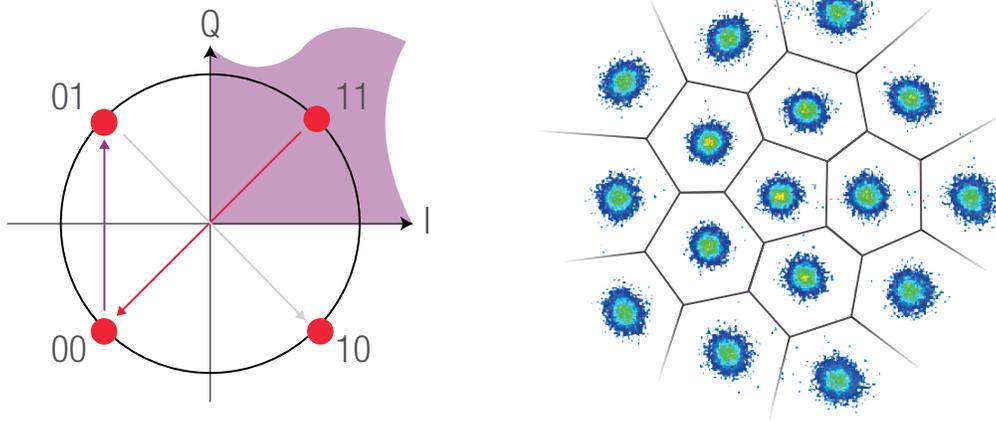


Figure 52. In QPSK, decision based on I and Q value, in more advanced formats on distance to nearest neighbor

Quality Rating of Coherent Measurements

Complex optical modulation schemes require new methods on the transmitter and on the receiver side. The previous chapter showed that coherent transmission is not in all aspects more complex than conventional OOK. Digital signal processing (DSP), as an integral part of coherent receivers provides relief from fundamental concerns about the impairing effects of dispersion. This benefit is the big plus of the coherent concept.

However, other distorting influences within the system need to be considered, and a thorough error analysis is mandatory. The natural question is: *Are new concepts needed here, too?*

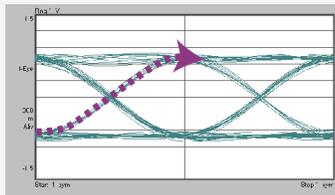
Traditional quality parameters

Let's look at the quality measures already known from OOK. There is the bit-error ratio (BER) and the Q-factor, which can be estimated from the eye diagram assuming a Gaussian noise distribution.

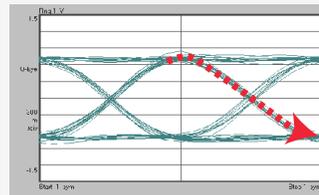
For the case of QPSK modulation, which is a widely used complex modulation format in 100-Gbps transmission systems, eye diagrams can be plotted to deduce these just mentioned quality parameters. Even though phase is the only varying parameter in QPSK, it is more practical to use two eye diagrams, where the I- and Q components are mapped to two independent eye diagrams.

In Figure 53, in the I eye diagram there is a transition from '0' to '1' denoted in purple, and for Q, in red a transition from '1' to '0'. As I and Q component are decoupled, this information cannot be projected to a symbol transition in the IQ-diagram without ambiguity. The left example could be a transition from '01' or from '00' to '11' or to '10', so four possible transitions are under consideration. The same can be noted when trying to map the Q-transitions.





Only I values versus time are visible



Only Q values versus time are visible

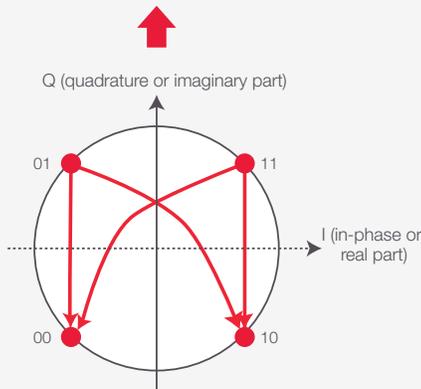
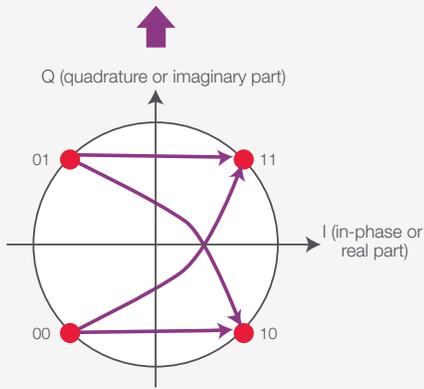


Figure 53. Eye diagrams for I and Q component of QPSK-signal (one polarization). A mapping to symbol transitions in the IQ diagram is always ambiguous. A timing skew between I and Q path is only visible in the IQ diagram

Such ambiguity may not necessarily create a problem. However, there is a distortion in the IQ plot that is not reflected in the two separate eye diagrams. In Figure 53, the bent diagonal transitions in both IQ plots point to a timing skew between the I and Q paths of the modulator. The I is obviously before the Q, which cannot be seen in the two eye diagrams.

For more advanced formats things get even more complicated. Look at Figure 54 with a special 16-QAM format. How would this be mapped to eye diagrams?

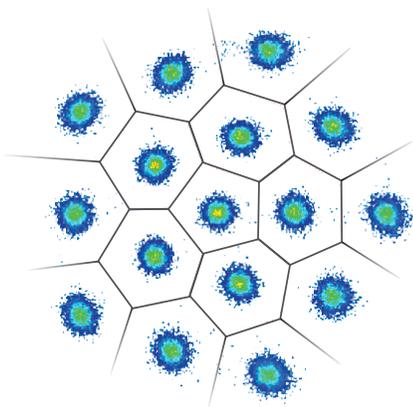


Figure 54. Measured constellation diagram of special 16-QAM format

The figure also reminds us that in complex modulation the decision on which bit sequence has been received is made in the IQ plane. Hence it makes much more sense to judge quality in the same picture. After all, some distortions have already proved to be more visible in this view.

Error vector magnitude (EVM)

Here again people learned from the RF community, where this problem was solved years ago with the intuitive approach of taking the distance of a measured point from the closest ideal constellation point. This concept leaves no ambiguity and works for any modulation format that can be represented in a constellation diagram.

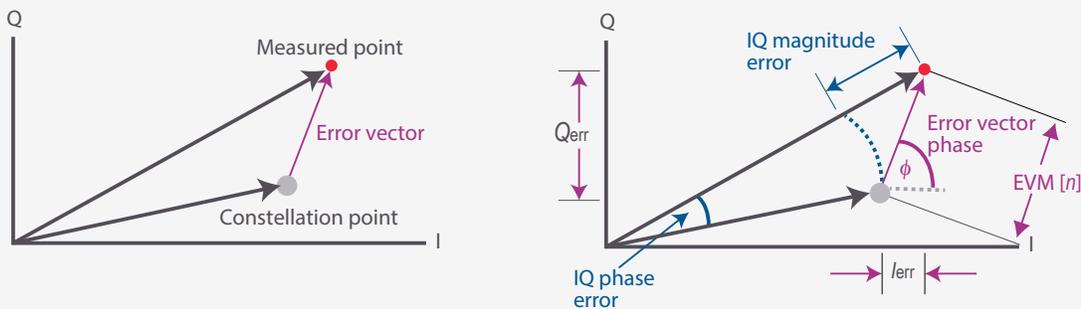


Figure 55. Definition of error vector and error vector magnitude

Figure 55 illustrates this for one measured point. The error vector magnitude $EVM[n]$ is the Euclidean distance between the measured point and the ideal reference point:

$$EVM [n] = \sqrt{I_{err}[n]^2 + Q_{err}[n]^2}$$

where n is the symbol index and $I_{err} = I_{Meas} - I_{Ref}$ and $Q_{err} = Q_{Meas} - Q_{Ref}$.

The normalized root mean square average EVM is defined as:

$$\%EVM = \frac{\frac{1}{N} \sum_{k=0}^{N-1} \sqrt{I_{err}[k]^2 + Q_{err}[k]^2}}{|\text{Peak reference vector}|}$$

where N is the number of measured points taken into account for calculating the EVM_{rms} (rms = root mean square).

The division by the magnitude of the peak reference vector serves for normalization. Other definitions use the average magnitude of all reference vectors or also the average power. This may be a trap when trying to compare EVM values.

In the example QPSK signal shown in Figure 56, the measured red dots around the four constellation points result in 5% error vector magnitude.

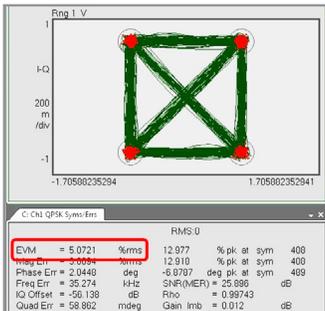


Figure 56. EVM measured on a QPSK signal

Signal to noise ratio (SNR)

From the EVM, the signal to noise ratio (SNR) can also be derived. It is also called modulation error ratio (MER) and defined as the ratio of average symbol power of the transmitted signal to noise power. This includes any term that causes the symbols to deviate from their ideal state:

$$\text{SNR} = 10 \log \frac{\sum_N (\text{IQ reference vector at symbols})^2}{\sum_N (\text{Error vector at symbols})^2}$$

BER estimation

For the case where there is only Gaussian noise, the BER can be predicted from the EVM (normalized to the average magnitude of all reference vectors) directly.

$$\text{BER (EVM)} = \frac{2 * \left(1 - \frac{1}{\sqrt{N_{\text{symbols}}}}\right) * \text{erfc} \left(\sqrt{\frac{3}{2 * (N_{\text{symbols}} - 1) * \text{EVM}^2}} \right)}{\log_2(N_{\text{symbols}})}$$

N_{symbols} is the number of symbols in the constellation diagram, *erfc* denotes the complementary error function and \log_2 is the logarithm with base 2. If the EVM does also contain contributions from other distortions, the prediction becomes more complicated. The measured BER does not increase as much as predicted if the simple model is used, hence the predicted BER provides an upper limit. This is quite obvious for the case where the signal does not contain any noise but the constellation is distorted. In that case EVM will have a certain value and hence the predicted BER will be non-zero; however, the actual number of bit errors may be zero.

Get a bigger picture from the EVM

So far, the EVM has only been investigated at the symbol times. Looking at the transitions in the IQ plots and plotting EVM versus time or frequency can help to find the root causes of distortions.

In Figure 57 there are two examples. On the left one, the measured signal is compared to a reference signal of “infinite” bandwidth; on the right side, the same measurement is compared to a reference signal created with a raised cosine filter. Both show the same EVM value at symbol times.

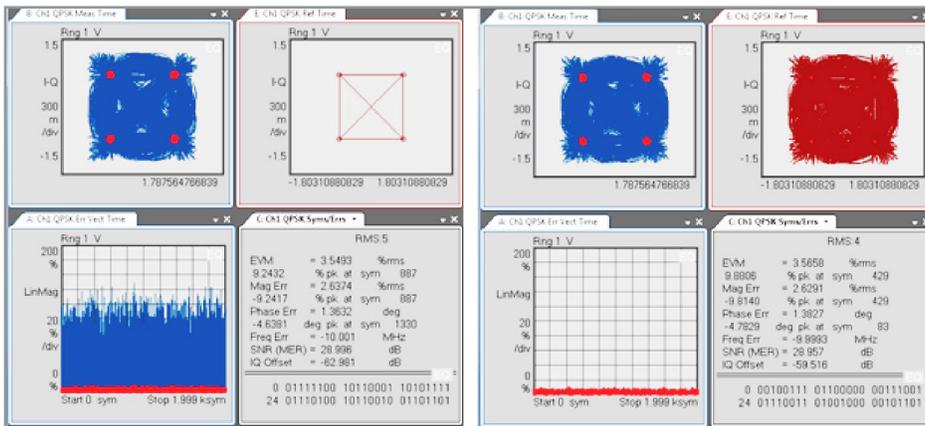


Figure 57. Same EVM at decision time but EVM vs. time reveals problems for “infinite” bandwidth (on the left side) but not for raised cosine filter (right side)

But looking at the EVM evolution over time (lower hand right windows respectively), the deviation of the measured signal from the “infinite” bandwidth reference signal reflected is reflected by high EVM values. If the raised cosine filter used in the transmitter has the same characteristics as the filter used for the reference signal, the EVM is low also during transition times between the symbols (right side). This analysis can be helpful to uncover unwanted behavior of your transmitter signal also during the transition phase from one symbol to the next,

Apart from the EVM, there are other error parameters deducible from the IQ plots that help us find the root causes of problems within the optical system.

Gain imbalance

$$\text{IQ gain imbalance} = 20 \cdot \log_{10} \left(\frac{I_{\text{magnitude}}}{Q_{\text{magnitude}}} \right)$$

The gain imbalance compares the amplitude of the I signal with the amplitude of the Q signal and is expressed in dB.

Figure 58 gives an example of a gain imbalance of about 2 dB, indicating a problem. The I and the Q magnitudes differ by a factor of 1.26.

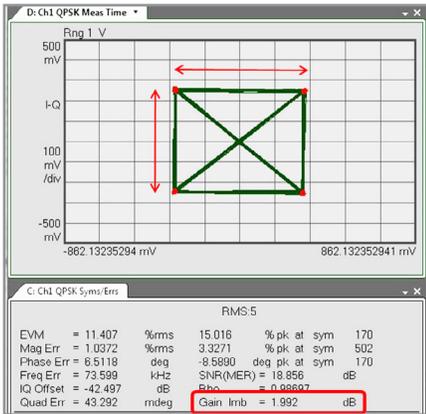


Figure 58. Gain imbalance: In the IQ plane, the I magnitude is bigger than the Q magnitude

The gain imbalance could be due to an unbalanced RF drive amplitude from the Mach-Zehnder modulator on the transmitter side.

IQ offset

IQ offset (also called IQ origin offset) describes the offset of the constellation diagram (Figure 59) from the origin (the ratio of the power at the center frequency to the average signal power).

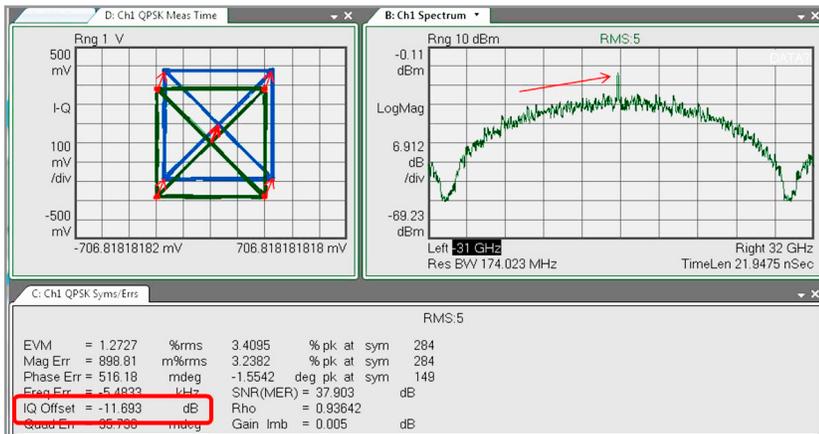


Figure 59. IQ offset: The IQ diagram is shifted away from the origin

This result indicates the magnitude of the carrier feed-through signal. When there is no carrier feed through, IQ offset is zero (-infinity dB). It is given as the ratio of signal to offset:

$$\text{IQ Offset} = 20 \cdot \log_{10} \left(\frac{\text{signal}}{\text{offset}} \right)$$

An IQ offset is often caused by a DC offset either in the I or Q path of the modulator or a small RF drive amplitude and a wrong bias point.

Quadrature error

The quadrature error quantifies the aberration from orthogonality between the I and Q. Ideally, I and Q should be orthogonal (90 degrees apart).

In the case in Figure 60, there is a quadrature error of almost 10 degrees, which means that I and Q are 80 degrees apart.

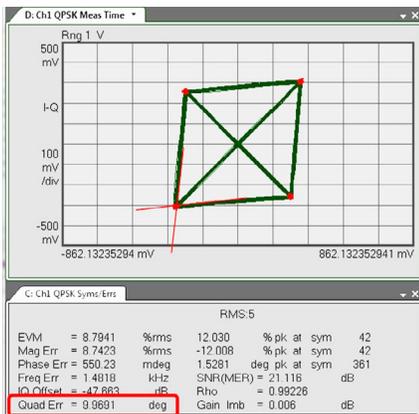


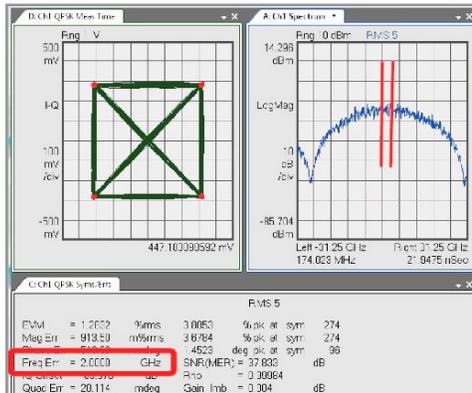
Figure 60. Quadrature Error: The I and Q phases are not orthogonal

A quadrature error usually points to a problem on the transmitter side, where the IQ 90° phase shifter may have the wrong bias point. In this case, also the eye diagram is distorted.

By adjusting the bias voltages of the Mach-Zehnder interferometer, orthogonality of the I and Q path can be achieved.

Frequency error

The frequency error shows the frequency difference between the carrier frequency and the local oscillator frequency. This error data is displayed in Hertz and is the amount of frequency shift that must be added in the digital domain to achieve carrier lock. The maximum frequency error that can be compensated is dependent on the used modulation format (compare Figure 61).



| Format | Maximum frequency error |
|----------|-------------------------|
| QPSK | 9.6% symbol rate |
| 16-QAM | 4.8% symbol rate |
| 32-QAM | 3.15% symbol rate |
| 64-QAM | 4.65% symbol rate |
| 128-QAM | 0.3% symbol rate |
| 256-QAM | 0.3% symbol rate |
| 512-QAM | 0.15% symbol rate |
| 1024-QAM | 0.15% symbol rate |

Figure 61. Example of frequency error and the maximum frequency error for different modulation formats. Note: the frequency error does not influence the error vector magnitude measurement

IQ magnitude error

The IQ magnitude error is the difference in amplitude between the measured signal and the ideal reference signal (compare Figures 55 and 62):

$$\text{IQ magnitude error} = \sqrt{I_{\text{Meas}}^2 + Q_{\text{Meas}}^2} - \sqrt{I_{\text{Ref}}^2 + Q_{\text{Ref}}^2}$$

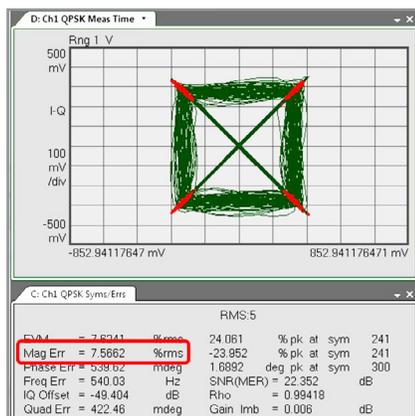


Figure 62. Example of magnitude error

Low frequency intensity noise from the transmitter laser, for example, can cause a magnitude error.

IQ phase error and laser linewidth

The phase error is the phase difference between the ideal IQ reference signal and the IQ measured signal measured at the symbol time (see Figures 55 and 63):

$$IQ \text{ phase error} = \arctan\left(\frac{Q_{\text{Meas}}}{I_{\text{Meas}}}\right) - \arctan\left(\frac{Q_{\text{Ref}}}{I_{\text{Ref}}}\right)$$

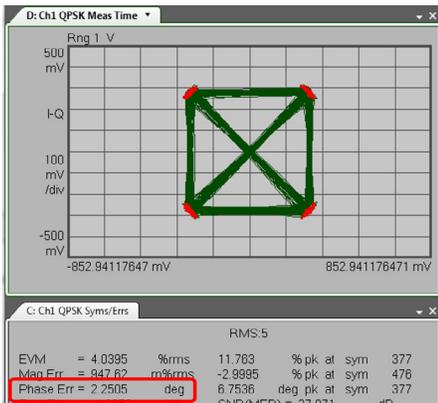


Figure 63. Example of a phase error

Phase error can be caused by phase noise from the carrier or local oscillator laser. It can also result in a time varying frequency error. The phase noise of a laser is typically quantified using the laser linewidth. It is possible to estimate the laser linewidth using the following procedure. Using the Kalman Filter Phase Tracking algorithm (see [Kalman Filter Based Estimation and Demodulation of Complex Signals](#)), the phase error can be evaluated over time and via Fourier transformation, the phase error spectrum can be obtained. By fitting a model to the phase error spectrum (see [Narrow Linewidth CW Laser Phase Noise Characterization Methods for Coherent Transmission System Applications](#)¹), the laser linewidth can be estimated as one of the free fitting parameters.

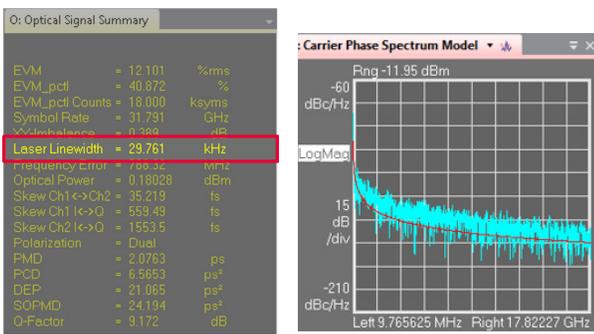


Figure 64. Laser linewidth estimation (left) and fit to carrier phase spectrum (right)

1. S. Camatel and V. Ferrero: Narrow Linewidth CW Laser Phase Noise Characterization Methods for Coherent Transmission System Applications, Journal of Lightwave Technology, Vol. 26, No. 17, September 1, 2008

IQ skew

The IQ skew measures the timing skew between I- and Q-signals of each polarization at the transmitter. This is done by measuring the phase difference of the symbol clock using the following equation:

$$\arctan(\alpha) = \left(\frac{\frac{1}{T} \int_0^T f(t) \cdot \sin(\omega t) dt}{\frac{1}{T} \int_0^T f(t) \cdot \cos(\omega t) dt} \right)$$

$\arctan(\alpha)$

The IQ skew will result in a distorted IQ plot and an increased EVM. When viewing the overlay of the respective I and Q-Eye-Diagrams, it can be seen that they are shifted with respect to each other. In case of very clean constellations, it can also be observed that the upwards and downwards 45° transitions will take different paths.

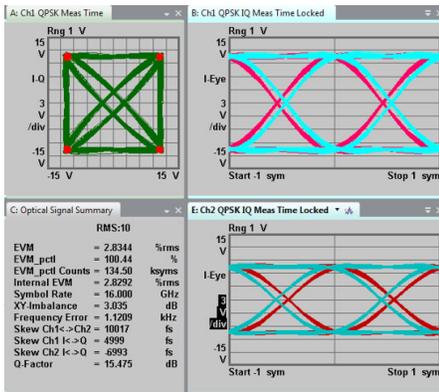


Figure 65. Example of IQ skew

X-Y polarization skew & imbalance

There can also be a timing skew between x and y polarization. It is calculated with the same equation as the IQ skew. The x-y polarization skew is not a critical parameter for signal quality measurements but actual network receivers might only tolerate a certain amount of x-y skew before the respective bit streams get out of sync. Note that the Keysight OMA software only reports a numeric value for the x-y skew. The respective eye diagrams will not show a skew due to the applied timing corrections.

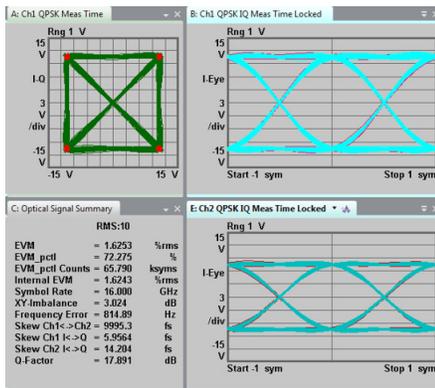


Figure 66. Example of imbalance skew

The x-y polarization imbalance is caused by a difference in power levels for the x- and for the y-polarization. The maximum power variation ΔP_{pol} of the optical power levels P_x and P_y calculates as follows:

$$\Delta P_{pol} = 10 \cdot \log_{10} \left(\frac{P_x}{P_y} \right)$$

The Future of Coherent Data Transmission

Complex optical modulation may be complex not only in a mathematical sense. But it is spectrally more efficient than any direct transmission format, which plays a role especially for longer distances and higher data rates. However, even in shorter distances such as metropolitan networks and data center interconnects (DCI), the traditional direct-detection transmission formats are partly abandoned for coherent modulation. Generally speaking, the higher the data rate, the more difficult it becomes to use direct-detection techniques even over short distances while on the other hand coherent transmission techniques can be simplified when optimizing them for short reach. This means that the cross-over point, where coherent detection becomes the technology of choice moves to shorter and shorter distances as the data rates increase. The future will tell for which applications coherent detection will catch up with direct detection in terms of price, size and power consumption.

Learn more at: www.keysight.com

For more information on Keysight Technologies' products, applications or services, please contact your local Keysight office. The complete list is available at: www.keysight.com/find/contactus

